CUK-H05316-15-9146596 ⑮

Economic Theory, International Trade and Quantitative Economics

Essays in Honour of Professor P.N. Roy

Edited by

Asis K. Banerjee
Professor of Economics
University of Calcutta

and

Biswajit Chatterjee
Reader in Economics
Jadavpur University

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

University of Calcutta
1996

© University of Calcutta
Senate House, Calcutta 700 073

Published by
Department of Economics
University of Calcutta

Printed in India by Pradip Kumar Ghosh
at the Calcutta University Press,
48, Hazra Road, Calcutta : 700 019

Price : Rs. 200

G. 146596

List of Contributors

1. Dr. Rajat Acharya — Department of Economics,Burdwan University, Burdwan, West Bengal, India.
2. Dr. Sarmila Banerjee — Department of Economics, Calcutta University, Calcutta, India.
3. Dr. Parantap Basu — Department of Economics, Fordlham University, Bronx, New York, USA.
4. Dr. Gautam Bhattacharya — Department of Economics, University of Kansas, Lawrence, Kansas, USA.
5. Dr. Tapan Biswas — Department of Economics, University of Hull, Hull, U.k.
6. Dr. Satya Chakravarty — Economic Research Unit, Indian Statistical Institute, Calcutta, India.
7. Prof. Biswajit Chatterjee — Department of Economics, Jadapur University, Calcutta, India.
8. Mr. Jaydip Dutta — Department of Economics, R. K. Mission College, Narendrapur, South 24 Parganas, West Bengal, India.
9. Dr. Amitabha Krishna Dutta — Department of Economics, University of Notre Dame, Notre Dame, Indiana, USA.
10. Dr. Dipak Ghosh — Department of Economics, University of Stirling, Stirling, Scotland, U.K.
11. Dr. Manash Ranjan Gupta — Department of Economics, Jadavpur University, Calcutta, India.
12. Dr. Subal Kumbhakar — Department of Economics, University of Texas, Austin, Texas, USA.
13. Dr. Kajal Lahiri — Department of Economics, State University of New York, Albany, New York, USA.

14. Dr. Sugata Marjit — Economic Research Unit, Indian Statistical Institute, Calcutta, India.

15. Dr. Bhaswar Moitra — Department of Economics, Jadavpur University, Calcutta, India.

16. Dr. Ishita Mukhopadhyay — Department of Economics, Calcutta University, Calcutta, India.

17. Dr. Rabindranath Mukhopadhyay — Department of Economics, Calcutta University, Calcutta, India.

18. Dr. S. Mansoob Murshed — Northern Ireland Economic Research Centre Belfast, Northern Ireland, U.K.

19. Dr. Nlandu Mamingi — International Economic Department, The World Bank, Washington D.C., USA.

20. Dr. Bibhas Saha — Indira Gandhi Institute of Development Research, Bombay, India.

21 Dr. Anjana Samanta (Bhattacharyya) — Centre for Urban Economic Studies Calcutta University, Calcutta, India.

22. Dr. Asani Sarkar — Department of Economics, University of Illinois at Urbana–Champaign, Urban, Illinois, USA.

23. Dr. Prabirjit Sarkar — Centre for Studies in Social Sciences, Calcutta, India.

24. Dr. Anindya Sen — Indira Gandhi Institute of Development Research, Bombay, India.

25. Dr. Somnath Sen — Department of Economics, University of Birmingham, Birmingham, U.K.

Contents

Introduction VII

Economic Theory

Interaction Between Formal and Informal Credit Markets in Agriculture : A Theoretical Analysis

Manash Ranjan Gupta 3

Saving Behaviour of The Traditional Sector and Development of A Dual Economy

Dipak Ghosh 16

Return Migration, Onward Migration and Local Job Search : A Sequential Approach to Migration

Gautam Bhattacharya 27

Monetary Uncertainty, Velocity and the Nominal Interest Rate : An Expository Note

Parantap Basu 68

Welfare Loss from Rent-Seeking for Budgetary Allocation

Satya R. Chakravarty 75

Weak* Axiom of Independence and the Non-Expected Utility Theory

Tapan Biswas 83

Share Tenancy and Family Resources

Jaydip Datta & Rabindranath Mukhopadhyay 94

Layoff Compensation and Entry Deterrence

Bibhas Saha & Anindya Sen 108

Non-Walrasian Static Exchange Problem

Ishita Mukhopadhyay 123

Joint Provision of Public Goods

Asani Sarkar 136

International Trade

On Keynesian Unemployment in a Fix-Flex Price Model of a Small Open Economy With Two Variable Inputs
Bhaswar Moitra 179

International Capital Movement, Trade Policy and Unemployment
Sugata Marjit & Rajat Acharya 204

International Debt, Terms of Trade Volantility and North-South Macroeconomics Realtions
S. Mansoob Murshed & Somnath Sen 216

Dumping and Trade Policy under The Condition of Demand Uncertainty and Technological Inflexibility
Sarmila Banerjee 240

Competitive South and Oligopolistic North: A Macro Model on Kaldorian Line
Prabirjit Sarkar 270

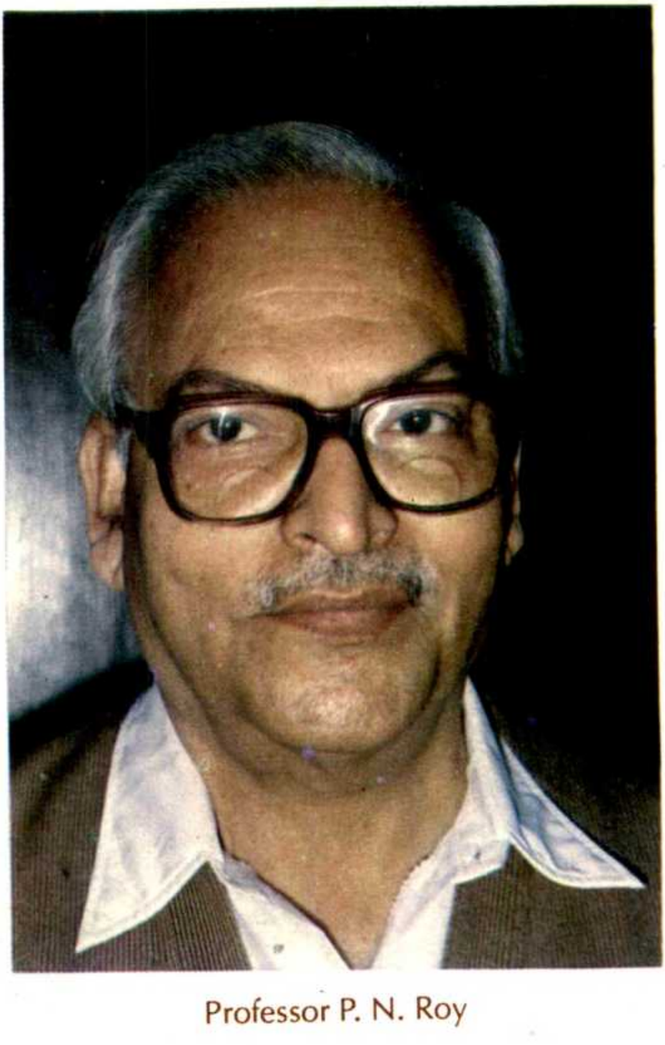
International Trade and Uneven Development: Implications of Some North-South Models
Amitava Krishna Dutt 279

Quantitative Economics

Granger Causality and Misspecified Vector Autoregressions
Kajal Lahiri & Nlandu Mamingi 315

Modelling Indirect Production with Expenditure Constraints: A Farm-level Study of Rice Production in West Bengal
Anjana Samanta (Bhattacharyya)
Subal C. Kumbhakar 330

Economics of Health, Nutrition and Development Theory and Evidence
Biswajit Chatterjee 356



Professor P. N. Roy

Introduction

Professor Prabuddhanath Roy belongs to the small class of economists who have a sustained interest both in Economics as an academic discipline and in the practical issues relating to public policy formulation. Within academic economics, again, his interest ranges over both economic theory and applied economics. So far as practical policy–making is concerned, the government has sought his assistance in a wide variety of matters, the most important among them being the formulation of the Five Year and the Annual Plans for the State of West Bengal. As a teacher of Economics, first in various undergraduate colleges in West Bengal, and then for a long time, in the postgraduate Department of Economics, Calcutta University, Professor Roy has inspired generation of students to think creatively about economic theory and policy, including international trade and quantitative techinques. Whenever any student looked for references for advanced reading materials in any branch of economics, he invariably would rush to Prabuddha-da for help. His reputation as an economicst and his pleasing personality have also won him many admirers and well-wishers among economists who are *not* formally his students.

In the present volume, we have collected a number of papers from among this large set of colleagues and students who wish to record their intellectual debt to Professor Roy, or their appreciation of his contributions in general on the occasion of his retirement from Calcutta University as a teacher of Economics. At present, he is the Pro-Vice Chancellor (Academic Affairs) at the University of Calcutta. Constraints of space forced us to be selective in the choice of the papers, but, on the whole, we believe, the papers in this volume reflect the main thematic interests of this respected teacher of economics.

The papers presented here deal broadly with three areas of Economics : economic theory (in the sense in which this term is usually understood in the profession), international trade, and quantitative economics. Within each of these areas the individual papers deal with a wide variety of topics. A degree of randomness, therefore, turned out to be unavoidable in the serial arrangement of the papers. The first group of papers is devoted to economic theory. Manas Ranjan Gupta's paper presents a theory of determination of the rate of interest on informal credit in the agricultural sector when there is a delay in the disbursement of formal credit, and the length of the delay varies inversely with the bribes from hte farmer and the lender of informal credit to the official granting formal credit. The effects of subsidising formal credit on the rate of interest on informal credit and the total cost of production of the farmers are analysed in this framework.

Dipak Ghosh attempts an analysis of the implication of *availability of savings* and the *form* of savings in the rural sector for industrialisation of a dual economy. It is shown that the latter is at least as important as the former in determining the overall growth rate of the economy.

One of the most important aspects of the process of economic development is migration of labour between the rural and the urban sectors. The existing literature on this topic, however, does not give sufficient attention to the question of how the individual decides whether or not to migrate. In the absence of a thorough analysis of such decisions many of the migration function proposed in the literature would appear to be *ad hoc* assumptions. Gautam Bhattacharyya seeks to fill this gap in the literature in his sequential model of job search and migration. He shows that under fairly weak assumptions there exist reservation sets of wages and revealed locational qualities which completely specify under what conditions an individual would search for a job locally, search at a new location ("onward migration"), search at a previous location ("return migration"), or not search at all.

One of the *stylised facts* in the recent history of monetary economics (at least as far as the advanced economies are concerned) is a fall in the income-velocity of circulation of money. This has given rise to a controversy as to whether this is to be interpreted as a contradiction of the monetarist position or a consequence of some changes in the mode of operation of the monetary authorities. Parantap Basu examines this controversy with the help of a parametric example of a cash-in-advance model a la Lucas and Stokey. In his model an increase in monetary uncertainty acts as a tax on the holding of "credit goods" which induces agents to switch from credit to cash goods, thus reducing velocity and the nominal rate of interest. The result is consistent with Friedman's monetary uncertainty hypothesis about the decline in velocity.

Satya Ranjan Chakravarty's paper suggests a welfare theoretic measure of rent-seeking induced by the government budget. The measure can be regarded as the proportion of the total spending which, for a given budget allocation, one would be prepared to sacrifice to eliminate all changes in the budget proportions allocated to different categories in a year over the preceding year.

An important area of research in economic theory is decision-making under risk. A basic result in this field is the Von Neumann-Morgenstern Expected Utility Theorem. However, some of the assumptions of this theorem have been called into question by empirical researchers. In particular, the "independence" axiom does not stand up very well to empirical tests. Tapan Biswas introduces a weaker assumption called weak* assumption of independence and shows that this assumption is strong enough to have important implications for decision-making.

Jaydeep Datta and Rabindranath Mukhopadhyay attempt to show how land lease decision of a (landless) household are influenced by its family endowments, particularly those which are not marketed or marketed imperfectly. They focus, in particular, on managerial skills and other human resources. The paper by Bibhas Saha and Anindya Sen deals with strategic behaviour of firms. It is shown that for strategic reasons a firm may benefit from the presence of strong labour union that demands lay-off compensation, or from regulations which require the firm to pay lay-off compensations.

Ishita Mukhopadhyay contributes a methodological paper. She considers the problem of compatibility between the economic logic (the subjective construct) implied by a theory and the applied mathematical techniques used to analyse the theory. She devises a simple compatibility test between the two and applies it in the context of a specific theory viz, the theory of non-Walrasian static exchange equilibrium.

The theory of public goods is a central topic in the theory of public finance. Asani Sarkar reanalyses the public goods question in a framework of relevance to an economy like India. Agents privately fund a public project of uncertain costs. It is shown that if the difference between the current and the expected future cost technology is *small*, the agents will initially behave as if the true cost is high for sure. Later on, however, the agents may decide to fund lower-cost alternative even prior to finding out whether that technology actually becomes available. Socially inefficient bargaining procedures which involve over-funding of the project may also occur.

Bhaswar Moitra studies Keynesian unemployment in a temporary equilibrium model of a small open economy with variable inputs, labour and capital. All prices except the rental rate are fixed *a priori* and the rental rate adjusts to clear the capital market. It is shown that Keynesian equilibria (i.e. equilibria where there is unemployment caused by demand failures) exist for particular configuration of the parametres of the system. The effects of changes in money-financed government expenditure on the endogenous variable are also analysed.

One of the important areas of theoretical research in international trade has been the North-South models of uneven development, which try to specify the working the advanced countries (North) in relation to backward labour surplus (South) economies in terms of alternative theoretical perceptions.

Sugata Marjit and Rajat Acharya use a neo-classical model of North-South trade relations with international capital mobility and domestic unemployment, and examines the effects of alternative tax/tariff policies on national welfare and unemployment in a typically backward economy. They demonstrate that the South, which has labour surplus and unemployment, imports northern capital and intermediate inputs to produce final output for the northern market exclusively, will benefit from a tax on foreign capital rather than a tariff on its importables, in terms of expansionary effects on employment. A set of sufficient conditions has been derived to ascertain this expansionary effect. The interesting result of the paper is that the condition which guarantees employment expansion as a consequence of tax on foreign capital, yields the opposite results when tariffs of final goods are imposed.

While the North-South trade relations have important implications for long-run growth equilibrium in the world economy, the most complex and fascinating aspect of this relation belongs to macroeconomic poicy interdependence among these two groups of countries. The implication of macroeconomic policy changes on the magnitude and dynamics of foreign debt is an important policy issue which merits theoretical scrutiny. S. Mansoob Murshed and Somenath Sen has constructed a monetary model of North-South relations where there is accumulated foreign debt of the South having balance of trade deficit with the North. They argue that in the medium term, the debt crisis of the Southern countries is the result of contractionary monetary policies in the North to contain inflation, thereby raising world interest rates and the burden of debt repayment accompanied by a sharp decline in Southern terms of trade. The asymmetry in adjustment mechanisms in the two groups of countries, as well as in the policy response were responsible for the sustainance of this crisis in the medium term, the authors argue.

One of the important area of concern in the realm of trade policy is the effects of dumping. This is an old practice, against which precarutionary measures are being taken in terms of specific legislations in different countries, advanced or backward. Sarmila Banerjee studied the effects of dumping in the presence of technological inflexibilities and demand uncertainties. Loss of domestic production, as well as reduction in consumer's surplus in the importing country have their implications on national welfare. Since the welfare loss is jointly caused by technological rigidities, and the power of the importing country to influence its terms of trade, she advocates a first best policy intervention in the form of either an import tariff or a production/consumption subsidy.

Prabirjit Sarkar adopts a more specific Kaldorian model to point out the implications of the difference in market structures in the North and the South on the long-run growth profiles of the two regions. He argues that South as a whole can not speed up its long-term growth rate through export promotion policies alone. Moreover, his model casts doubt on the Prebisch–Singer hypothesis of a secular decline in the terms of trade of the South in terms of a mere rise in the mark-up rate in the manufacturing sector of the North-perhaps one needs to look to other channels for a better and more comprehensive explanations.

Amitava Krishna Dutt also is one of the leading contributors in this line of research. His paper in this volume addresses the process of uneven development in terms of three types of analytical models to highlight the importance of sector-specific increasing returns as the source of uneven development, the effect of technical progress in the North on its dependence on Southrn primary intermediate goods, and the possible future effects of the growth of newly industrialised nations on the international asymmetries.

The section on quantitative economics contains three papers. Kajal Lahiri and Nlandu Mamingi developed a method of Granger Causality and Vector Autoregressions to demonstrate the role of the degree of cointegration of variables in the distortions caused by misspecification errors. Such errors and distortions are likely to diminish larger the sample size, but the rate of disappearance of distortions depends crucially on the high or weak cointegration of variables. The authors recommend the use of ECM method instead of simple VAR even in large samples.

The paper by Subal Kumbhakar and Anjana Bhattacharya (Samanta) uses a translog version of indirect production function in the context of rice cultivation in West Bengal. Technical characteristics of such production functions apart, their paper estimates the elasticities of output with respect to different inputs for small and large farms engaged in rice cultivation in the state of West Bengal. In addition to the well known constraints like small holding size and paucity of capital equipment, the limited availability of working capital is also identified to be an important factor limiting growth. This has important implications for policy reorientation towards making formal sector loans more accessible to small and marginal farmers in a resource-constrained agrarian economy like that of India.

The last paper in this volume is on nutrition and poverty in West Bengal written by Biswajit Chatterjee. An attempt is made in this paper to measure the extent of nutritional deprivation in West Bengal on the basis of secondary information, particularly the NSS consumer expenditure surveys. Chatterjee's construction of simple indices of IDI and FDI signifying malnutrition and hunger bears simple but forceful insights to the living conditions of the people of West Bengal over almost two decades. The analysis is aggregative, but it is suggestive of an empirical methodology to encompass disaggregated household based comparisons as well. His conjecture that the social sector development in India as well as in West Bengal is likely to receive secondary importance in government's policy in an era of liberalization and global competitiveness, brings into focus an important issue in the recent literature on theory of public action and the political economy of hunger and deprivation in backward nations.

The themes of the papers are of contemporary relevance in the advancing literature on economic theory, international trade and applied quantitative economics, and it is hoped that the contributions come up to the high academic standards set by Professor Roy as a teacher of economics and a scholar par excellence. We sincerely hope that this collection will be useful to the serious scholars in the subject.

Asis Banerjee
Biswajit Chatterjee

Calcutta, August 1996

Economic Theory

1

Interaction Between Formal and Informal Credit Markets in Agriculture : A Theoretical Analysis

Manash Ranjan Gupta

1. Introduction

For financing the cost of cultivation,the farmers in Indian agriculture find access to two types of credit: (i) informal credit available from the rural money-lenders at high rates of interest; and (ii) formal credit available from the nationalized banks and co-operatives at subsidized interest-rates. Various theories have been developed to explain the determination of the rate of interest on informal credit' But none of these considers the role of interaction between the formal credit market and the informal credit market on the determination of interest-rate on informal credit.

In the presence of subsidized formal credit, the farmers go to the informal credit market if either the formal credit is inadequate in supply or it is not available in time. Recent empirical studies focus on the second problem. Sarap (1990B), analysing the village-level data of the Sambalpur District of Orissa, India, has shown that the large farmers have more or less timely availability of formal credit, while the small farmers are denied. His study shows that it takes the small and marginal farmers more than a month to obtain the formal credit. Even after the loan has been sanctioned the small and the marginal farmers have to wait for a few weeks to collect it. Among the factors responsible for this delay, emphasis has been given on the patronage, arbitrariness and corrupt practices pursued by the officials of the formal sector credit institutions due to an inter-play of political and social factors. In addition to Sarap (1990), this aspect has been pointed out by Ladman and Tinnermeir (1981), Adams and Vogel (1986), Ghatak (1977, 1983), Benjamin (1981), Lele (1981) etc.

In this paper, we consider the interest-rate determination on the informal credit when there is delay in disbursing the formal credit and the delay is the result of the corrupt practices of the official of the formal credit agency. If the formal credit is not available at the beginning of the crop-cycle, the farmer takes informal credit for that interval. The delay is determined by the behaviour of the official who takes bribe from the farmer and the money-lender. The farmer bribes the official to lower the delay; and the money-lender bribes the official to increase the delay and hence to raise the farmer's demand for informal credit. So the rate of interest on the informal credit that the money-lender determines having a monopolistic advantage in that market is not independent of the delay of disbursement of formal credit and the corrupt behaviour of the Official. Subsidized formal credit given to the farmers can not bring its full benefit in this case because a part of the subsidy is pocketed by the official. For example, a reduction in the rate of interest on formal credit raises the rate of interest on informal credit and the cost of the farmer.

2. The Model

The agricultural crop-cycle is described by the time-horizon [O, T] and there is a given time-distribution of cost per acre of land over the cycle. Let C (t) be the cost per acre of land at the time point t; and for the sake of simplicity, it is assumed that C (t) = C (constant)[2] for all $t \in [O, T]$. Formal credit is available at the time point $T^* \in [O, T]$; and hence CT^* is the farmer's requirement of informal credit per acre of land.

There are three economic agents in the system :

(i) the farmer who takes loan; (ii) the money-lender who lends informal credit and determines its interest-rate; and (iii) the official whose behaviour determines the delay in disbursing the formal credit, T*, and who is used to take bribe from the other two agents.

2.1 The Official

While analysing the behaviour of the official, we start with the following functions :

$$U = U(Y_F, Y_M, 1\text{-}L) \qquad \text{..................(1)}$$

This is the utility function of the official and it satisfies all the standard neo-classical properties. Y_F and Y_M are the amounts of bribe he takes from the farmer and the money-lender respectively; and both the bribes are consumed. However, these two bribes are not perfectly substitutes. L is his working-time and 1-L is his leisure time.

$$T^* = G(L) \text{ with } G'(.) < 0 \qquad \text{..................(2)}$$

The equation (2) implies that the disbursement of formal credit is speeded up if the official works more.

$$V = V(T^*, \lambda) \text{ with } V_i > 0 \text{ for } i = 1, 2 \qquad \text{..................(3)}$$

Is the disutility function of the official. Bribe taking is an illegal activity and the official will be punished if he is caught. Let λ be an index of penalty (may be an amount to be paid as fine) imposed on the official if he is caught. Also a delayed disbursement of credit is considered as inefficiency in the service. So the disutility of the official is a positive function of the delay in the credit disbursement and of the index on penalty.

The official will maximize his net welfare (U-V) through the choice of L; and the first-order condition of maximization is given by the following :

$$-U_3 \cdot V_1 \cdot G'(\cdot) = 0 \qquad \text{..................(4)}$$

where U_3 is the marginal utility of the third argument, leisure. This means that, in equilibrium, the marginal net welfare of leisure is zero.

The total differential of equation (4) is given by the following:

$$[\, U_{33} - V_{11} (G'(\cdot))^2 - V_1 G''(\cdot) \,] \, dL - V_{12} \, G'(\cdot) \, d\lambda$$
$$= U_{31} \, dY_F + U_{32} \, dY_M \qquad \text{..................(5)}$$

Note that $U_{33} < 0$ because this implies the diminishing marginal utility of leisure. Now we assume the following :

(i) $V_{11} > 0$ and (ii) $G''(\,.\,) > 0$. Assumption (i) implies that the marginal disutility of T* is an increasing function of T* ; and assumption (ii) implies a concave production function where output is the speeding up of credit disbursement and the official's working time is the only input.

Suppose that the bracketed term in the L.H.S. of (5) is denoted by A. So A is negative. Hence,

$(dL / dY_F) = (U_{31} / A)$; and

$(dL / dY_M) = (U_{32} / A)$.

Using equation (2) and these two expressions, we have

$(dT^* /dY_F) = G'(.) (U_{31} /A)$;
$(dT^* /dY_M) = G'(.) (U_{32} /A)$; and
$(dT^* /d\lambda) = [G'(.)]^2 (V_{12} /A)$.

We assume V_{12} to be positive. This implies that the marginal disutility of delay is a positive function of λ. Hence,

$(dT^*/ d\lambda) < 0$.

because both A and G' (.) are negative. Sign of (dT^*/dY_F) and (dT/dY_M) depend on the sign of U_{31} and U_{32}. We assume that $U_{31} < O$ and $U_{32} > 0$. Hence, $(dT^* /dY_F) <0$ and $(dT^*/dY_M) > 0$. This leads to the following proposition :

PROPOSITION 1 : T^* is a negative function of Y_F and λ, and a positive function of Y_M.

So we have the following function derived from the behaviour of the official :

$T^* = T^* (Y_F, Y_M, \lambda)$ ....(6)

with $(\delta T^* / \delta Y_F) < 0$; $(\delta T^*/ \delta Y_M) > 0$; and $(\delta T^*/ \delta\lambda) < 0$.

Note that $U_{31} < 0$ and $U_{32} > 0$ imply that the marginal utility of leisure of the official is a negative function of the amount of bribe given by the farmer and a positive function of the bribe given by the money-lender.

2.2 THE FARMER

Let r and i be the rates of interest on informal credit and formal credit available to the representative farmer. So in the presence of delayed formal credit and bribing, the additional production cost, Z_F, of the farmer is given by

$Z_F = (r-i)\ CT^* + Y_F$

where $x = (r-i)\ C$ is the additional interest cost per unit of delayed time.

The farmer minimizes this additional cost through the choice of Y_F. The first-order condition of minimization is

$x. (\delta T^* / \delta Y_F) = - 1$ ....(7)

and its total differential is given by the following :

$$(\delta T^* / \delta Y_F)\, dx + x. [(\delta^2 T^* / \delta Y^2_F)\, dY_F + (\delta^2 T^* / \delta Y_F \delta\lambda)\, d\lambda + (\delta^2 T^* / \delta Y_F\, \delta Y_M)\, dY_M] = 0 \qquad \text{..................(8)}$$

Here $(\delta^2 T^* / \delta Y^2_F) > 0$ is the second-order condition of minimization ; and for the sake of simplicity, we assume that $(\delta^2 T^* / \delta Y_F\, \delta Y_M) > 0$ and $(\delta^2 T^* / \delta Y_F\, \delta\lambda) > 0$.

$$\text{Hence, } (dY_F / dX) = -\frac{(\delta T^* / \delta Y_F)}{x\,(\delta^2 T^* / \delta Y_F^2)} > 0 ;$$

$$(dY_F / dY_M) = -\frac{(\delta^2 T^* / \delta Y_F\, \delta Y_M)}{(\delta^2 T^* / \delta Y_F^2)} < 0 ; \text{ and}$$

$$(dY_F / d\lambda) = -\frac{(\delta^2 T^* / \delta Y_F\, \delta\lambda)}{(\delta^2 T^* / \delta Y_F^2)} < 0.$$

So the optimum bribe function of the farmer is given by the following :

$$Y_F = Y_F\,(x, Y_M, \lambda) \qquad \text{..................(9)}$$

Hence, Y_F is a positive function of X and a negative function of Y_M and λ.

Given Y_F , the farmer faces a lower cost with a higher X or λ ; or a lower Y_M . The marginal net gain of bribing of the farmer $[1 - x\,(\delta T^* / \delta Y_F)]$ rises when the rate of interest in the informal credit market and /or the penalty on bribing increases or the money-lender's bribe falls. So the farmer bribes more in all these cases.

2.3 THE MONEY-LENDER

Let g be the interest-rate at which the money-lender can get fund. It may be the interest-rate in the urban credit market. We assume that g > i. Subsidized formal credit is available only to the farmers. Government of India has adopted such a policy. The professional money-lenders can not get this subsidized credit[3]. Also the farmer can not get fund at the rate of interest g. This is because the land or agricultural output may not be used as colateral in the urban credit market and the urban lenders may not be willing to lend the farmers directly due to various risks and uncertainties associated with agricultural production. But the rural money-lender generally keeps a close contact with the urban credit market and can get the credit there.

The net-interest income of the money-lender is given by the following :

$$Z_M = (r - g) \, . \, C.T^* - Y_M$$

or, $$Z_M = x.T + (i - g) \, C.T^* - Y_M \qquad \text{................(10);}$$

and this is maximized with respect to r and Y_M.

The first-order conditions of maximization are given by the followings :

$$T^* + [x + (i - g) \, C]. \, (\delta T^* / \delta Y_F). \, Y'_F \, (x) = 0 \qquad \text{..................(11);}$$

and $$[x + (i - g) \, C] \; [\, (\delta T^* / \delta Y_M) + (\delta T^* / \delta Y_F) \, (\delta Y_F / \delta Y_M)] = 1 \qquad \text{..................(12).}$$

Solving these two equations, one can obtain the equilibrium values of r and Y_M.

2.4 CREDIT SUBSIDY POLICY

In this subsection we point out the problem associated with the policy of reducing the rate of interest on formal credit.

Equation (11) can be expressed in the following form :

$$(x + (i - g) \, C) \, \beta \, \theta = x \qquad \text{..................(11.1)}$$

where $\beta = - \, (\delta T^* / \delta Y_F) \, (Y_F / T^*) > 0$;

and $\theta = (Y'_F \, (x) \, x / Y_F) > 0$.

Here β is the elasticity of delay in disbursement of formal credit with respect to the bribe given by the farmer ; and θ is the elasticity of farmer's supply of bribe with respect to his additional interest cost per unit of delayed time. We assume both these two elasticities to be constant. Here βθ plays the role of interest-elasticity of demand for informal credit.

Now from (11.1) we have,

$$\beta\theta, dx + \beta\theta \, . \, C. \, di = dx \; ; \text{ and, } dx = C \, (dr - di)$$

Hence, we have,

$(\beta\theta - 1)(dr - di) + \beta\theta . di = 0$ or, $(\beta\theta - 1)\, dr = - di$

We assume that $\beta\theta > 1$. This is because the price - elasticity of demand is to be greater than unity in the case of equilibrium of a monopolist. Hence, we find

$$(dr / di) = - \frac{1}{\beta\theta - 1} < 0 \qquad \text{..................(13).}$$

This leads to the following proposition :

PROPOSITION 2: A fall in the rate of interest on formal credit leads to an increase in the rate of interest on informal credit.

The intuitive explanation of the proposition is the following : The equation (11.1) can be expressed as follows :

$$r - i = (g - i)\,[\,\beta\theta / (\beta\theta - 1)]$$

Here (r - i) is the effective price of informal credit ; (g - i) is its effective unit (marginal) cost ; and $\beta\theta$ is equivalent to the interest-elasticity of the demand for informal credit. So the equation (11.1) is the equilibrium condition of the monopolist money-lender. The monopolist, facing a rise in marginal cost, must raise the price at a higher rate because the price-elasticity of demand is greater than unity. So a reduction in i raises (g - i) and hence leads to a higher rate of increase in (r - i) because $\beta\theta > 1$. So r should rise with a fall in i.

This result is important in the context of available empirical findings. It was thought that the credit subsidy policy would weaken the position of the money-lenders and lower the rate of interest on informal credit. But the empirical study of Sarap (1990A) shows that the rate of interest charged by the informal money-lenders particularly to the poor borrowers is still very high[4].

Note that, equation (12) can also be expressed in this form:

$$[x + (i - g)C]\,(T^* / Y_M)\,(\gamma - \beta\alpha) = 1 \qquad \text{..................(12.1).}$$

Here $\gamma = (\delta T^* / \delta Y_M)\,(Y_M / T^*) > 0$;

and $\alpha = (\delta Y_F / \delta Y_M)\,(Y_M / Y_F) < 0$.

Here γ is the elasticity of delay in disbursement of formal credit with respect to the bribe given by the money-lender; and α is the elasticity of farmer's bribe supply with respect to the bribe given by the money-lender. We assume γ and α to be constant.

Note that,

$(T^* / Y_M) = (1/\gamma).(\delta T^* / \delta Y_M).$

and, hence, from equation (12.1), we have,

$[X + (i - g) C] (1/\gamma). (\delta T^* / \delta Y_M) (\gamma - \beta\alpha) = 1$

or, $(X / \beta\theta) (1/\gamma).(\gamma - \beta\alpha) (\delta T^* / \delta Y_M) = 1$ [using (11.1)]

or, $M.X (\delta T^* / \delta Y_M) = 1$ ....(12.2)

where, $M = (1/\beta\theta)(1/\gamma)(\gamma - \beta\alpha) > 0.$

Taking the total differential of (12.2) ; we have

$$\left[(\delta T^*/\delta Y_M) + \left(\frac{\delta^2 T^*}{\delta Y_M\, \delta Y_F}\right).(\delta Y_F/\delta X)\, X\right] dX$$

$$= - X.(\delta^2 T^*/\delta Y_M^2)\, dY_M$$

or, $(\delta T^*/\delta Y_M)(1 + \rho\theta)\, dx = -x. (\delta^2 T^*/\delta Y_M^2)\, dY_M$

or, $(1+\rho\theta)(dx/x) = \pi(dY_M/Y_M)$ ....(14).

Here $\rho = (\delta^2 T^* / \delta Y_M\, \delta Y_F)[Y_F/(\delta T^*/\delta Y_M)] > 0$; and

$\pi = -(\delta^2 T^*/\delta Y_M^2)[Y_M/(\delta T^*/\delta Y_M)] > 0.$

then $dx > 0$ implies $dY_M > 0.$

Also $(dx/di) = [(dr/di) - 1] C < 0.$ [$\because$ $(dr/di) < 0$].

So $(dY_M/di) < 0$. This leads to the following proposition :

PROPOSITION 3: With a reduction in the rate of interest on formal credit, the money-lender gives more bribe to the official.

The intuitive explanation of this proposition is the following: A reduction in the rate of interest on formal credit raises the farmer's interest cost per unit of delay, x. So the farmer bribes more and this raises the marginal income of bribe of the money-lender. So the money-lender will bribe more in the new equilibrium.

Now we want to consider the effect of credit subsidy policy on the additional cost of the farmer. We know that,

$Z_F = x.T^* + Y_F$

Hence,

$$(\delta Z_F / \delta i) = (\delta x / \delta i)T^* + (\delta Z_F / \delta Y_F) \left[(\delta Y_F / \delta x)(\delta x / \delta i) + (\delta Y_F / \delta Y_M).(\delta Y_M / \delta i)\right] + x.(\delta T^* / \delta Y_M)(\delta Y_M / \delta i) \quad ......(15)$$

Here $(\delta x / \delta i) < 0$; $(\delta Z_F / \delta Y_F) = 0$ in the equilibrium of the farmer ; $(\delta T^* / \delta Y_M) > 0$ and $(\delta Y_M / \delta i) < 0$.

Hence, $(\delta Z_F / \delta i) < 0$.

So we have the following proposition :

PROPOSITION 4: A fall in the rate on interest on the formal credit leads to a rise in the additional interest cost of the farmer.

The intuition behind the proposition is the following :

With a reduction in the rate of interest on formal credit, the money-lender on the one hand charges higher interest-rate on informal credit and on the other hand gives more bribe to the official (which in turn increases the delay of disbursement of formal credit). So the cost of the farmer is increased.

This last proposition is most important to the analysis. This signifies that the farmer may not receive the full benefit of the credit-subsidization policy in the presence of delayed disbursement of formal credit and the corrupt behaviour of the official of the formal credit agency.

2.5 EFFECTS OF PENALTY ON BRIBING:

In this section we consider the effect of a change in penalty on bribing. From equation (11.1), it is clear that a change in λ does not affect x and hence r. But using equation (12.2), we can show the effect of its change on Y_M. We have,

$$\left[(\delta T^* / \delta Y_M \, \delta Y_F)(\delta Y_F / \delta Y_M) + (\delta^2 T^* / \delta Y_M^2)\right] dY_M + (\delta^2 T^* / \delta Y_M \, \delta Y_F).(\delta Y_F / \delta \lambda)\, d\lambda = 0$$

Hence, $(dY_M/d\lambda) < 0$ because $(\delta^2 T^*/\delta Y_M\, \delta Y_F) > 0$;

$(\delta Y_F/\delta Y_M) < 0$; $(\delta Y_F/\delta \lambda) < 0$ and $(\delta^2\, T^*/\delta Y_M^2) < 0$.

This leads to the following proposition :

PROPOSITION 5 : An increase in the penalty on bribing leads to a lower bribe of the money-lender.

Next, we consider its effect on the cost of the farmer.

$$\frac{dZ_F}{d\lambda} = x\left[\left(\frac{\delta T^*}{\delta Y_M}\right)\left(\frac{dY_M}{d\lambda}\right)+\left(\frac{d\,T^*}{d\lambda}\right)\right] < 0.$$

$$\left[\because \left(\frac{dx}{d\lambda}\right) = 0;\ \&\ \left(\frac{dZ\,F}{dYF}\right) = 0\right]$$

So we can prove the following :

PROPOSITION 6 : An increase in the penalty on bribing lowers the additional cost of production of the farmer.

The intuitive explanations of these two propositions are the followings : As the penalty on bribing is increased, the official lowers the delay given his bribe-income. This lowers the cost of the farmer at the same level of bribes given by him and the money-lender. So in the first-stage, the money-lender's bribe being unchanged, the farmer bribes less. This produces a negative effect on the marginal gain of bribing of the money-lender. But the marginal gain of bribing of the money-lender is reduced, as he bribes more. So in the new equilibrium, he will bribe less. The increased penalty on bribe-taking then lowers the delay of credit-disbursement directly as well as indirectly. So the additional cost of the farmer is reduced.

2.6 UNCERTAINTY :

The model developed in Section 2 is based on the assumption that the farmer and the money-lender have full information about the bribe of the other agent. But the problem is that giving or taking bribe is illegal ; and hence this full-information assumption seems to be inappropriate. A better way of looking into the case seems to assume that the farmer, while deriving the optimal value of Y_F, should find Y_M to be stochastic.

Suppose that Y_M is a random variable. Then T* is also a random variable and let its expected value and variance be E (T*) and V. Then expected value of Z_F and its variance are given by $E(Z_F) = (r - i)\,C.E\,(T^*) + Y_F$; and $\mathrm{Var}\,(Z_F) = (r - i)^2\,C^2.V$

We assume that the representative farmer maximizes utility defined as a negative function of expected extra-cost, $E(Z_F)$, and a function of its variance. The utility is a negative (positive) function of the variance if the farmer is risk-averter (risk-lover). The utility is independent of the variance if the farmer is risk-neutral.

The analsis in this section is identical to that in the previous section under the assumption that T* is a separable function of Y_F and Y_M. In this case, variance of T* is independent of Y_F. So the maximization of utility with respect to Y_F is identical to minimization of the expected extra-cost, $E(Z_F)$; and this leads to the first-order condition given by equation (7).

3. Conclusion :

The paper presents a theoretical analysis of the rate of interest determination in the informal credit market in a backward agriculture where there is provision of subsidized formal credit and the supply of formal credit is constrained by its delay in disbursement and the corrupt behaviour of the official of the formal-credit agency. The main problem that the analysis points out is that a policy of subsidized formal credit can not bring its entire expected benefits to the farmers. For example, it raises the rate of interest in the informal credit market and also aggravates the extent of corruption in the form of increased bribe-income of the official. This also raises the additional cost of the farmer. So in the presence of corrupt behaviour of the formal sector lender, it becomes important to reconsider the desirability of the policy of supplying subsidized formal credit to the farmers. Obviously, the increased legal actions and penalties partly solve the problem. But the implementation of all these involves a lot of administrative problems and the Indian experience in this context is not very hopeful.

REFERENCES

Adams, D.W and Vogel, R C., 1986, 'Rural financial markets in low income countries Recent controversies and lessons', *World Development*, Vol 14, 477-87.

Basu, K , 1984, '*The less developed economy A critique of contemporary theory*', Oxford, Basil Blackwell.

Bardhan, P.K., 1983, *Land, labour and rural poverty*, Oxford University Press

Benjamin, M.P., 1981, *Investment projects in agriculture ; principles and case studies* ; London, Longman

Bhaduri, A., 1977, 'On the formation of usurious interest rates in backward agriculture', *Cambridge Journal of Economics*, Vol. 1, 341-352.

Bottomley, A., 1975, 'Interest-rate determination in underdeveloped rural areas', *American Journal of Agricultural Economics*, Vol 57

Ghatak, S., 1975, 'Rural Interest rates in the Indian Ecomomy', *Journal of Development Studies*, Vol. 11, 190-201.

Ghatak, S., 1977, 'Rural Credit and the cost of borrowing : Interstate variations in India', Journal of Development Studies, Vol. 13, 102-124.

Ghatak, S., 1983, 'On Interregional variations in Rural Interest rates in India', *Journal of Developing Areas*, Vol. 18, 21-34.

Gupta, M.R., 1987, 'A nutrition based theory of interlinkage', *Journal of Quantitative Economics*, Vol.3, 189-202.

Ladman, J.R. and D.W. Adams, 1978, 'The rural poor and recent performance of formal rural financial markets in the Dominican Republic', *Canadian Journal of Agricultural Economics*, Vol. 26, 43-50

Ladman, J.R. and Tinnermeier, R.L., 1981, 'The political economy of agricultural credit : the case of Bolivia', *American Journal of Agricultural Economics*, Vol. 63, 66-72.

Lele, U., 1981, 'Cooperatives and the poor : A comparative perspective; *World Development* , Vol. 9, 55-72.

Sarap, K. 1987, 'Transactions in rural credit markets in Western Orissa, India', *Journal of Peasant Studies*, Vol. 15, 83-107.

Sarap, K., 1990A, 'Interest rate determination in backward agriculture : The role of economic and extraeconomic control', *Cambridge Journal of Economics*, Vol 14, 93-108.

Sarap, K , 1990B, 'Factors affecting small farmers' access to institutional credit in Rural Orissa, India', *Development and change,* Vol. 21, 281-307

FOOT NOTES

[*] This is a revised version of an earlier draft. I am indebted to an anonymous referee for his valuable comments Remaining errors are solely mine.

1 See Bottomley (1975), Bhaduri (1977), Basu (1984), Gupta (1987) etc

2. It is a simplifying assumption.

3. However, in many cases, the large farmers themselves play the role of money-lender in the informal credit market and they can get subsidized formal credit. In this paper we consider a money-lender whose alternative activities do not include farming.

4 See Sarap (1990A) ; Table 2.

Savings Behaviour of The Traditional Sector and Development of A Dual Economy

Dipak Ghosh

0. Introduction

An understanding of the savings behaviour of the rural sector is crucial for any researcher interested in the process of economic development of labour surplus, dual economies, be it from the point of view of empirical, policy oriented investigations, or that of purely theoretical analyses. [See, for example, Preobrazhensky (1925), Dobb (1966) in the context of Soviet economic planning, Lewis (1954) and the theoretical literature on development of dual economies that followed. See articles in Adelman and Thorbeck (1966), the publication of Lewis, Kalecki (1971, 1976) to mention but a few]. Professor Prabuddha Nath Roy, in his role of an economic planner and policy maker as an active Member of the State Planning Board of the State of West Bengal, has to deal with the issues related to inter sectoral linkages from the point of view of empirical application. A glimpse of his view on the theoretical aspect of the linkage between the two sectors can be found in the precise survey of "Optimal Investment and Trade Policies in a Dual Economy" in Roy (1986).

This article attempts to take a look at the implications of *availability of savings* and *the form in which the savings is available* in the rural sector, on the industrialisation or modernisation of dual economy. At the very outset, it is to be made clear that the spirit in which this article is being written is in accordance with the Indian tradition that it is not possible to repay *guru reen.* This is just a very small attempt to acknowledge a very large *reen* which will continue to remain unsettled.

The two sectors being addressed to in this article are denoted as *traditional* and *modern.* In the standard dual economy, post - Lewis (1954), literature while the terms, traditional, agricultural and rural sectors are used interchangeably to denote what Lewis described as traditional sector, terms such as modern, industrial and urban have been used to denote Lewis's *modern* sector.

Though some degree of overlapping of these terms are unavoidable in any discussion of development of a dual economy, one should not lose sight of the fact that orginally the difference between the traditional and modern sectors was defined in terms of the differences of *technology* used in these two sectors. The traditional sector which uses 'primitive' technology is assumed to have a *viable*, albeit low, standard of living. In Lewisian world agricultural activities which use 'modern' technology will belong to the modern sector and so for that matter will a rural firm which uses 'advanced' technology.

In what follows, in Section I, we will try to take a look at the savings behaviour in a non (or partially) monetized traditional sector and the implication of such behaviour on the process of economic development. In Section 2, we will take a look at a scenario where the traditional sector is monetized. The process of economic development here is defined in the sense of Lewis --- capital accumulation and migration of labour from land based to non-land based activities.

1. Non-Monetized Traditional Sector

In the majority of developing countries we come across a kind of duality between the traditional and modern sectors, not only in terms of technology in use but also in terms of their degree of monetization. As the process of economic development starts, the traditional sector is predominantly non-monetized. What is the definition of a monetized sector? One can define a monetized sector very simply as a sector where money plays all its roles., i.e., where it is used as medium of exchange, a unit of account and a means of savings (store). A purely non-monetized sector is either a *pure subsistence sector*, where agents produce just enough to satisfy their own needs, or a sector where goods or services are *bartered*. However, unlike in the monetized sector where the agents try to maximize possible pruduction surplus to take advantage of the market, in the non-monetized sector the agents do not operate in the interest of the market [see Chandavarkar (1977) for alternative definitions of monetization]. In this sense, monetization is synonymous with *commercialization*. Commercialization of a sector is measured by the proportion of outpur sold or *marketed surplus* in total output. We now make an attempt to examine the savings behaviour in a non-monetized sector of an economy. Within the sector to be investigated, we concentrate entirely on the behaviour of small and medium food grains producing peasant farmers as farmers producing commercial crops and the 'big' farmers who produce food grains as commercial crops are already monetized.

In an economy where the food grains producers are monetized, the total output produced by a farmer over and above the consumption need of the farmer and his family, which is the *marketable surplus*, is sold and money obtained in exchange is either saved or invested after the purchase of goods and services the farmer himself does not produce. Of course the farmer may choose not to sell all or part of his surplus if he is convinced that the price of his produce is going to rise in future. In this case the farmer's stock holding behaviour is purely speculative. When the price of a commodity is high stocks are unloaded in the market there by moderating the price rise. Alternatively when the price is low, stocks are held back which helps price from falling any further. In a monetized economy the supply of marketed surplus is positively related to price.

The savings behaviour in a non-monetized peasant sector is completely different from that described in the last paragraph. In a non-monetized sector the farmers save, when they can, in kind rather than in cash. The nature of uncertainty under which peasant farmers have to operate is the reason behind this kind of savings behaviour. In a non-monetized, peasant economy the sale of output by the farmers is mostly 'distress sale'. The farmers having a target demand for cash, to buy the 'other sector's' goods, pay rent etc. When the price of his produce, food grains, is high a farmer can meet his target demand for cash by selling a relatively small portion of his produce. On the other hand, when the price of food grains is low the farmer has to sell a larger quantity to meet the given demand for cash. Thus the marketed surplus, for food grains, in a non-monetized sector is negatively related to price. Of course, when a farmer's income is high, due to high price of his produce, his ability to save is also high. However, in this sector, when a farmer can afford to save, he saves in kind. [See Mathur and Ezekiel (1961)]. What prompts this kind of savings behaviour is the attitude of the peasant farmers towards the risks and uncertainties under which they have to operate. We must understand the nature of such risks and uncertainties. For example, if there is a crop failure the farmer is compelled to borrow, to protect his family from starvation, from the available sources for lending, who are either big landlords or moneylenders. This kind of loan is usually in kind. As a peasant farmer has very little or no formal collateral, he may have to offer the small amount of land he owns as collateral or pledge promise of future labours (including those of his wife and children) during harvesting season to the landlord or the borrower may lose control over choice of crops in his own plot of land for the future. Given that the cost of a future crop failure may imply destitution for the farmer and his family, the farmers in this situation follow a strategy for *survival* rather than optimization. When they can save, they save in kind.[1]

What is the implication of such savings behaviour on the process of economic development? Let us try to sketch a simple algebraic model to investigate this. Let the numerals 1 and 2 denote the two sectors traditional and modern., $\mathbf{q}_{ij}$ the demand for the jth sector's output by the ith sector. Price of food grains, which is produced by the traditional sector alone, $\mathbf{P}_1$, is expressed in terms of the modern (manufacturing, industrial) sector's output. In other words $\mathbf{P}_2$ is assumed to be unity. Let $\mathbf{Y}_i$ denote the income of the ith sector, and $\mathbf{O}_f$ the total output of food grains. Let us consider only one period where the output of food grains is fixed. We then have :

$$Y_1 = P_1 O_f \quad (1) \qquad O_f = \bar{O}_f \quad (2)$$

The traditional sector's own demand for food grains can be expressed in the following form:

$$q_{11} = F(P_1, Y_2) \quad (3)$$

which gives us

$$q_{11} = f(P_1) \quad (4)$$

The industrial sector's demand for food grains can be written as

$$q_{21} = g(P_1, Y_2) \quad (5)$$

The total demand for food grains plus any stock in a given period must equal the total supply. This gives us the balance equation (6) bellow.

$$f\ (P_1) + g\ (P_1, Y_2) + S\ (P_1) = \overline{O}, \qquad (6)$$

where S (P_1) denotes total stock.

Differentiating (4) and (5) totally and differentiating the balance equation (6) with respect to $\mathbf{P}_1$ and after some messy manipulations one gets

$$\frac{dP_1}{dY_2} = \frac{\dfrac{-\delta g}{\delta Y_2}}{\dfrac{\delta g}{\delta P_1} + \dfrac{\delta f}{\delta P_1} + \dfrac{\delta S}{\delta P_1}} \qquad (7)$$

We can now denote industrial sector's income elasticity of demand for food grains as

$$\eta_{Y_2} = \frac{\delta g}{\delta Y_2} \cdot \frac{Y_2}{q_{21}},$$

the industrial sector's price elasticity of demand for food grains as

$$-\eta^2_{P_1} = \frac{\delta g}{\delta P_1} \cdot \frac{P_1}{q_{21}}$$

and the food grain producing sector's price elasticity of demand for food as

$$\eta^f_{P_1} = \frac{\delta f}{\delta P_1} \cdot \frac{P_1}{q_{11}} \qquad (8)$$

It should be pointed out that the last elasticity term is composed of both income and price elasticities of demand remembering equation (1). For example, when $\mathbf{P}_1$ decreases food grains become cheaper to the farmers giving rise to an increase in the farmer's demand for food. However, a decline in $\mathbf{P}_1$ implies a reduction in farmer's income and hence, his demand for the same. So what is the ultimate sign of this elasticity?

Writing the farmer's demand equation in full to include the price of sector **2**'s output, we get

$$q_{11} = D\ (P_1, P_2, Y_1)$$

We will now refer to a theorem on elasticities due to Wold and Jureen (1953). The theorem tells us that for an individual demand function for any *non-inferior* commodity the sum of own price elasticity, cross price elasticities and income elasticity is equal to zero. In symbol we, therefore, have :

$$\eta_{P_1} + \eta_{P_2} + \eta_Y = 0$$

Now, the first term in the equation above is the own price elasticity and hence has negative sign. The second term is the cross price elasticity of demand for food grains. As farmers have a target demand for the other sector's goods, as $\mathbf{P_2}$ increases farmers' demand for food grains must decrease making the cross price elasticity term negative. The third term, the income elasticity of demand is, of course, positive in sign. Expressing the elasticity relation in the following manner :

$$\eta_{P_1} + \eta_Y = -\eta_{P_2}$$

it is obvious that the sign of the sum of own price elasticity and income elasticity of demand for food grains is positve. Thus the sign of the combined income and food price elasticity in equation (8) is positive.

We now define γ as elasticity of S with respect to P_1. Given our discussion on the savings behaviour in the non-monetized sector this elasticity is of positive sign.

We can now re-write equation (7), using the elasticity symbols as:

$$\frac{Y_2}{P_1}\frac{dP_1}{dY_2} = \frac{\eta_{Y_2}}{\eta^2_{P_1} - \eta^f_{P_1}\frac{q_{11}}{q_{21}} - \gamma\frac{S}{q_{21}}} \qquad (9)$$

Equation (9) gives us the elasticity of food grains prices with respect to income in the modern sector. The assumption that the absolute value of the slope of the demand function for food in the modern sector (for a given level of Y_2) is greater than the slope of the food producing sector's demand curve and the stock function's slope combined leads us to conclude that the sign of the elasticity term in equation (9) is positive.[2]

G-146596

Let us now turn our attention to marketed surplus. We write equation (6) as

$q_{11} + q_{21} + S_F + S_T = \overline{O}_f$

where

$S = S_F + S_T$.

S_F and S_T denote farmers and traders stock respectively. Remember that the total demand for food in the modern sector, both for consumption and stock building by traders in a closed economy, can only be met by surplus marketed by the farmers. Denoting marketed surplus by MS we have the following relation based on equation (6).

$$q_{11} + MS + S_F = \overline{O}_f$$

Defferentiating this relationship above with respect to price of food grain P_1 and translating the resulting equation in terms of appropriate elasticities we obtain the crucial relationship as follows

$$\eta^{f}_{P_1}\frac{q_{11}}{P_1} + \eta^{M}_{P_1}\frac{MS}{P_1} + \gamma\frac{S_F}{P_1} = 0 \qquad (10)$$

We have already established that the combined food price and the farmers' income elasticity of demand for food $\eta^{f}_{P_1}$ is positive as is γ Hence $\eta^{M}_{P_1}$, the elasticity of marketed surplus with respect to price of food grains must be negative.

Equations (9) and (10) tell our story. The first of these equations tells us that in an economy a process of industrialization puts upward pressure on price of food grains and the second one tells us, if the food producing sector is non-monetized, then as price of food grains increases less of it will be marketed in the modern sector. In developing countries where food grains constitute the major portion, if not almost the entire, basket for *wage goods*, the inflationary implications of our analysis in such an economy is obvious. Thus during the process of development, *the form in which* savings are held becomes important.

2. Monetized Traditional Sector

Is monetization of the traditional sector the panacea that all labour surplus dual economies should aim to achieve? The answer to this question is, as it happens to be almost always in economics, that monetization is a necessary but not a sufficient condition that guarantees removal of all constraints to industrialization. The question of *availability of savings* becomes relevant here. For this part of the discussion we will carry out within the confine of the post-Keynesian, mainly Kaldorian, frame work. Our starting point is, once again, Lewis (1954) where Sector 1 is the only sector which produces and supplies food grains (wage goods). Sector 2 produces 'industrial goods'.

The most crucial point here is that the prices of food grains and those of industrial goods are determined under different regimes. Food grains prices are *flex price* i.e., prices which are demand determined. This is to say, that in the short to medium run, as the supply of food grains is relatively inelastic, their prices are determined by the height of the demand curve for food grains. On the other hand, the prices of the industrial goods are *fix price* or are prices which are predominantly cost determined. The, by now familiar, price relationship in a fix price frame work is given by

$$P_2 = (1 + \tau)\, wl \qquad (11)$$

where, $\mathbf{P_2}$ is the price of the industrial sector's goods in terms of food grains, τ is the profit mark-up, **w** is the unit labour cost in terms of food grains and **l** is the labour requirement per unit of outpur.[3]

Now, unlike in an economy where the traditional sector is non monetized, in the present scenario the supply of wage goods from the traditional sector to the modern sector will depend upon the amount of cash savings that the traditional sector can generate by selling their marketable surplus to the modern sector along with the former sector's demand for the latter sector's produce. If for, say, structural constraints like lack of capital, land,adequate infrastructure etc. the traditional sector's capacity to produce food grains gets restricted this is reflected on this sector's ability to save, and hence supply food grains to the modern sector. In the face of the inelastic supply of food grains any attempt at expanding economic activities in the modern sector can only give rise to inflationary pressure. This is because an attempt to expand activities in the modern sector will increase demand for labour and hence that of wage goods giving rise to an upward shift of the demand function for food grains. In a flex price regime, according to which prices of food grains are determined, this implies a rise in the prices of food grains. Food grains being the most crucial component of wage goods basket, this will give rise to an increase in **w** in equation (11) generating pressure of inflation in the economy as a whole.

The inflationary implication of lack of availability of savings from the food grains producing sector is but only one part of a whole picture. Let us assume in the line of Kaldor (1977) that the food grains producing sector is in equilibrium in the sense that they can sell all they want to at the ruling price. Also savings in this sector is translated into investment. However, assume further that the ability to produce for this sector is restricted due to some kind of infrastructural bottleneck. This then translates to : the amount of savings this sector can generate is also limited. Looking at the problem from the point of view of the modern sector, constrained savings in the traditional sector implies limited effective demand for this sector's goods from the traditional sector. Under a situation such as this, even if the modern sector itself is not constrained by the availability of savings lack of demand or expected demand for this sector's output will result in investors reducing their planned investment in the future. In this sense even though the traditional sector is in equilibrium the modern sector is not as in this sector planned investment is less than available savings. Incidentally, this possibility of planned investment being restricted by lack of effective demand is the intellectual basis of the doctrine of the foreign trade multiplier in the post-Keynesian growth literature.

It is to be noted that though a strategy of export led growth may alleviate the problem of lack of effective demand for the modern sector's output, so long as the modern sector is dependent upon the traditional sector for supply of wage goods, any constraint on Sector 1's ability to save will give rise to general inflationary pressure via food price inflation. In this sense expansion of domestic economic activities as a whole may co-exist with the pressure of inflation in developing countries.

3. Conclusion

In this short article we have tried to take a look at the importance of the traditional sector, the sector which is responsible for the provision of wage goods, in the process of growth and development of the modern sector in developing countries. The political economy of a society will determine the degree of monetization of the taditional sector there. The physical constraints on production are determined by the socio-economic-political conditions. A lack of appreciation of these issues and treating the problem of economic planning only as an optimizing problem can only give rise to frustrations.

Foot Notes

1. On the nature of risks and savings beaviour in developing countries see Ghosh (1986a)

2. If this condition does not hold then no stable price level will ever be achieved.

3. Fix price does not imply constant prices, it just means prices do not respond to fluctuations in demand as these fluctuations are moderated by stock adjustments. See Ghosh (1986b)

REFERENCES

Adelman I and Thorbeck E (eds.) (1966) : *The Theory and Design of Economic Development*, Johns Hopkins

Chandavarkar A G (1977) : "Monetization of Developing Countries" *IMF Staff Paper*, Vol. 24

Dobb M (1960) . *An Essay on Economic Growth and Planning*, Routledge & Keegan Paul

Ghosh D (1986a) : "Monetary Dualism in Developing Countries" *Economies et Societes*, Vol. 20

Ghosh D (1986b) · "Fix Price-Flex Price in Development Economics" *Australian Economic Papers*, Vol. 25

Kaldor N (1977) . "Equilibrium Theory and Growth Theory" in M Boskin (ed) *Economics and Human Welfare : Essays in Honour of Tibor Scitovsky*, Academic Press

Kalecki M (1971) · *Selected Essays on Dynamics of Capitalist Economy* 1933-1970, CUP

Kalcki M (1976) : *Essays on Developing Economies*, Harvester Press

Lewis W A (1954) . "Economic Development with Unlimited Supplies of Labour" *Manchester School*, Vol 20

Mathur P N and Ezekiel H (1964) : "Marketable Surplus of Food and Price Fluctuations in a Developing Economy" Kyklos, Vol 17

Preobrazhensky E A (1925) . "Economic Notes I On Goods Famine" in *The Crisis of Soviet Industrialization*, D A Filtzer (ed) Macmillan 1980

Roy P N (1986) : *Internatioanal Trade Theory and Practice*, Wiley Eastern

Wold H and Jureen L (1953) : *Demand Analysis*, Wiley

Return Migration, Onward Migration and Local Job Search : A Sequential Approach to Migration

Gautam Bhattacharya

1. Introduction

The objective of this paper is to analyze different aspects of the migration decisions of an individual. Typically, migration decisions are a part of a lifetime job search process across locations where the individual considers job opportunities at different locations, "quality of life" given by desirable and undesirable characteristics of different locations, and the nonpecuniary characteristics of different jobs. On the other hand, he also considers search costs associated with jobs and moving costs associated with locations; further, there may be quitting costs associated with jobs and monetary and nonmonetary costs associated with moving away from a location after residing there for a long period of time. The individual also considers possibilities of improvements in his own skills and in his own personal wealth while making his search/migration decisions.

Finally, the information about quality of different locations may be incomplete, and learning about the quality of a location after moving there can play an important role in the individual's decision about whether to continue staying there, or to migrate to a new location, or return to a previously visited location. Although a thorough analysis of an individual's search/migration decisions that incorporates all these elements appears to be a formidable task, it is clear that a satisfactory answers to the issues related to labor migration can only be obtained from a sequential model of search and migration where incomplete information about quality of locations and the job search process are considered explicitly.

The current literature on migration lacks any substantive analytic framework on the theory of individual migration. Most of the papers concentrate mainly on the empirical aspects of migration, which are very important for investigation of the critical factors behind return and onward migration, of the importance of information on migrants' decisions, of the nature of interregional migration flows in different countries etc.[see, Allen (1979), Davanzo (1983), Herzog and Schlottman (1983), Morrison and Davanzo (1987), Schlottman and Herzog (1982), Herzog, Hofler and Schlottman (1985), etc. for these and related issues]. The only recent paper that deals with the search/migration problem in detail is McCall and McCall (1987) which uses the powerful multi-armed-bandit model of statistical decision - making literature to show how the individual's migration decisions are affected by incomplete information about locations and job opportunities. In my paper, I have developed a model of search and migration of an individual to characterize his lifetime migration /search decisions under incomplete information about wage offers and quality levels at different locations.In this model, the problem of the individual is to decide from a menu of given locations with unknown quality (or attractiveness) , which location to choose for job search and eventual migration if an acceptable job offer is obtained. The true quality of the location is revealed to him after he migrates to a location and works there for one period. At that point, he has the options of - (a) staying at the same job at the same location, (b) choosing another new location for another round of job search and possible migration (this is denoted as onward migration in the literature), (c) looking for a job in the same location again at a lower search cost but with a known quality level (the option of local search) and (d) looking for a job at a previously visited location with known quality (the option of return migration). My model is analyzed by using a search- theoretic framework that uses the tools of dynamic programming and that is closer to the structure of papers on search under incomplete information [Burdett (1978), Wilde (1979, 1980), Lippman and McCall (1981) etc.]. Although this is a more conventional approach than the multi-armed bandit model employed by McCall and McCall (1987), my model appears to be more general in several respects because I have considered continuous distributions of wage and quality levels and considered the option of local search that is ignored in their model. Further, as shown in the concluding section of this paper, my model appears to have a more reasonable view of return migration and has a large number of empirically testable hypotheses that are absent from their model.

The basic conclusions of the model can be summarized as follows. I have shown that under rather mild assumptions, there exist reservation sets of wage and revealed location qualities that completely specify under what conditions the individual will search locally, or search at a new location for onward migration, search at a previous location for return migration or not search at all in every stage of the search / migration process. Also, I can answer questions about parameter changes, i.e. how local and national search costs, distribution of wages and location quality etc. affect the search and migration decisions in every stage. It is shown that return migration, when optimal, does not depend on the characteristics of the current location, whereas onward migration decisions are shown to depend on both the characteristics of the destination location and the known quality and wage levels of the current region as well. Further, the individual is shown to choose an optimal sequence of migration decisions where he will choose onward migration successively, terminating the migration process whenever he chooses the local search, no search or return migration option. Finally, because of the explicit consideration of the local search option, I have shown how search costs and wage distributions at other locations will affect the individual's job quit decisions in the current location and his onward and return migration decisions as well. These results constitute a new framework to analyze the process of return migration, onward migration and the relations between migration and local job turnovers.

These results will probably be of considerable interest to researchers pursuing empirical work in the areas of migration flows, of job turnovers, and of the determinants of onward vs. return migration. Further applications of these results (after necessary modifications) are possible in the fields of international migration from third-world countries and the brain drain (see for example Katz and Stark (1984, 1987)), and the phenomenon of return migration of residents belonging to one cultural/ethnic group who return to their native country after working at different places and possibly in different countries of the world (this phenomenon has been observed in detail by sociologists - see for instance King (1986)). Finally, the theoretical framework developed in this paper can be useful in the analysis of other search problems with incomplete information-for example, the consumer-purchase of

CALCUTTA UNIVERSITY CENTRAL LIBRARY

durable goods models (see Hey and McKenna (1981)), the exploration for natural resource problems, and the problem of allocation of resources for innovative activities (similar to Weitzman) (1979).

2. A model of onward migration, return migration and local job search.

There are I locations denoted by $i = 1, 2, 3, ..., I$ and the distribution of job offers in every region is known to the individual. The job offers in region i, for instance, are given by a probability distribution F_i (w) with a density function f_i (w), when w is a value of the wage associated with the offer. Although the individual knows the distribution of job offers for every region, he does not know about the set of characteristics (denoted in this paper as quality) that cannot be observed unless he migrates to a particular location by accepting a job there. We assume that these characteristics can be summarized by a variable q that measures their total pecuniary value to the individual. The individual then assumes (from whatever a priori knowledge that he may have) that the distribution of q in a location is given by a probability distribution with known density function a (q). Note that the individual's a priori knowledge is not assumed to be location specific, thus a (q) is the same for all locations.

The search and migration problem of the individual can be described in several stages. In the first stage, the individual is away from all the I locations and he decides to choose one location where he would search for a job first. I assume that the individual can search for a job at only one location at a time. The search cost in location i is given by c_i and search is assumed to be instantaneous. If a job offer at location i at wage w_i is accepted by the individual, then he actually moves to that location (I have not considered any moving costs here, although this could be introduced without changing the nature of the model). Thereafter, he works in that job for one period, receives w, and finds the value of the non wage (quality) characteristics of the location i. Now the individual enters the second stage of the migration / search process with three options.

The first option is to search for a job at any one of the remaining I-1 locations and migrate there in the event of a successful serarch. If the individual chooses this option, he effectively goes back to the first stage of the decision process, although the number of remaining locations to be searched is one less than the number of locations in the first round of national search. The second option is to search for another job in location i. This type of search, called local job search, has the following features. The individual looks for a job at location i by drawing offers from the wage distribution with the density function f_i (w). Search cost per search in location i is assumed to be L, which is presumably lower than all the c_i's, because the individual is searching from inside the location. The most important feature of local search is that the individual can assure himself of the known quality level q_i if he searches for a job locally at location i. Thus although local search reduces options of the individual by permitting him to search in the current location only, he can avoid the quality uncertainty that is associated with the national job search process. The third option for the individual is to stay at the current job at the current location,earn a wage of w and enjoy a quality level of w_i for ever. I assume that in this case his per period payoff is $q_i.w_i$ and thus the current value of his total lifetime income is $q_i.w_i /(1 - \beta)$, when β is the discount factor (the multiplicative form of the payoff function is chosen for its simplicity, other forms may also be used with similar results).

If the second option of local search or the third option of staying at the same job at the same location is chosen, then it is obvious that the individual has no more interest in migration and the search / migration process ends there, at least until any other parameters faced by the individual changes. However, if the first option mentioned above is chosen,then we have the case of onward migration where the individual moves to another location in the event of a successful search (say location j). Now, as the quality level of the location j, q_j, is revealed to him after he works there for one period at wage w_j, he can again choose between several options. The first three options are the same as above - he can migrate onward , search locally, or decide to stay at the same job at location j. However,there is another option available at this stage, that of return migration. Under this option, he can look for a job at location i where he held a job before. I assume, of course, that the job he quit before at wage w_i cannot be recalled, but the individual has the advantage about location i that he knows the true quality level of location i to be q_i. Thus, by looking for a job at a previously visited location i, he can be certain of the quality level q_i. The search costs for return migration are presumably different from those of local search (denoted by L here for all locations). Although it is clear that searching for a job from outside the location is more costly than searching for a job from inside the same location, it is not obvious whether the fact that the individual has visited the location before will affect his search costs for that location if he is currently searching for a job from outside that location. In any case, to avoid more complications, I assume that the search costs for searching from outside is always c_i for location i, irrespective of whether the individual has visited that location before. If he decides to accept a job in the previously visited location. i, then he migrates back to the location i and presumably ends the migration / search process.

The search / migration process of the individual, then, can be intuitively summarized as follows. After k rounds of migration to k distinct locations have already occurred, define N^k as the set of previously visited locations and the current location i, and M^k be the set of the locations not visited yet. In all periods of this discrete time model, $N^k \cup M^k = (1,2,3, ...,I)$. In period zero, N^o is the null set and M^o is the set of all I locations. In period one,after one round of migration, N^1 has just one element, the current location, and M^1 is the set of (I-1) locations i.e. all locations except the current location. In general, after k rounds of onward migration and with no return migration, N^k is the set of k locations including the current location, and M^k is the set of the remaining I- k locations. If at any time the option of local search is chosen the migration process terminates. Because, as the future is discounted, if it is not optimal to explore any other options at present, it will not be optimal to do so later.For the same reason, if at any time the option of return migration is chosen, the migration process is terminated.

The individual's objective is to maximize the expected value of his discounted life time income -

$$\sum_{t=1}^{\infty} \beta^{t-1}(w_t q_t - S_t) \qquad (1)$$

when w_t is his wage rate in period t (not in location t) , q_t is the value of the quality of the location in period t, and s_t is the amount spent on search in period t.

The expected value in (1) will depend on the stopping rule adopted by the individual at different stages of the search/migration process. As mentioned before, the individual can search in only one location in each period although search is instantaneous and he can select any locations that he likes. Thus, in the first stage, when he is searching for a job from outside all the locations, his decision about which location to choose for job search and eventual migration will depend on which location, if chosen, has the maximum expected total lifetime income.Let $\underline{W}_{ii}$ be the optimal stopping rule for location i in the first stage in the first round of migration i.e. if the individual has chosen to search for a job at location i and if he gets an offer of w_i, then,

if $w_i \in \underline{W}_{ii}$, he accepts the job and migrates to location i, if $w_i \notin \underline{W}_{ii}$, he continues search at location i.

Let $V^o(i)$ be the individual's expected value of lifetime income from following an optimal search/migration policy during all periods and starting from location i in the first stage., and let V^o be the expected value of his overall search/migration process if he follows an optimal strategy throughout. Then

$$V^o = \max [V^o(1), V^o(2), \ldots\ldots\ldots, V^o(I)] \qquad (2)$$

To evaluate $V^o(i)$, I will use the optimality principle of dynamic programming to claim

$$V^o(i) = -c_i + V^o(i)\,[1 - \int_{\underline{W}_{ii}} f_i(w)\,dw] + \int_{\underline{W}_{ii}} J_i(o,w)\, f_i(w)\,dw \qquad (3)$$

when the function J_i (o, w) is the optimal expected total payoff to the individual if his current location is i, and an offer of w is accepted in region i.

As mentioned before, after the worker accepts a job at wage w at location i, he moves to location i and starts working on the new job. He earns a wage of w in the current period and finds the true value of q in location i, thereby getting a payoff of q.w in the current period. Define the function H_i (1, w, q) as the expected total payoff after one round of migration has occurred and the worker has worked in location i for one period, has found the true value of q and has obtained a one period payoff of q.w.. Therefore,

$$J_i(o,w) = \int_q [qw + \beta H_i(1,w,q)\ a\ (q)\ dq] \tag{4}$$

To evaluate the expected payoff function H_i (1, . , .) after one round of migration, notice that the options in this stage are to stay at the same job at the same location, search locally for a new job at the same location with the assured quality level q, or start from the first stage again by looking for a job in one of the remaining I – 1 locations (the option of return migration does not exist in the first period).

1) If the individual decides to stay, his total lifetime payoff is equal to q.w / (1 – β).

2) If he decides to search locally, and adopts the optimal local stopping rule $\underline{W}_{Li}(\underline{q})$, his expected total payoff is defined to be V_L (i,q) when, as in equation (3)

$$V_L(i,q) = -L + \int_{\underline{w}_{Li}(\underline{q})} \{qw/(1-\beta)\}\ f_i(w)\,dw\ +$$

$$(1 - \int_{\underline{w}_{Li}(\underline{q})} f_i(w)\,dw)\ V_L(i,q) \tag{5}$$

3) If the individual decides to search at other locations for a job (i.e. at one of the other I – 1 locations with unknown quality), then I define V^1 as the optimal expected payoff from his search/migration process after one round of migration, and V^1 (j), j ≠ i as his expected payoff if he searches at location j after one round of migration and continues with an optimal strategy,

Thus H_i (1, w, q) will be given by

$$H_i\,(1, w, q) = \max\,[qw/\,(1-\beta),\ V_L\,(i, q),\ V^o\,(j)\ ;\ j \neq i] \tag{6}$$

To complete the description of my search/migration model, I need to specify how the optimal expected payoff after k – 1 rounds of onward migration, defined as V^k, is evaluated for any $k = 0, 1, 2, \ldots$. As before, let $V^k(j)$ be the optimal expected payoff if the individual decides to look for a job in location j after k–1 rounds of onward migration have already occurred. Also, as before, define $J_j(k, w)$ as the expected return from location j after an offer of w has been accepted and the worker moves to location j on the kth round of onward migration. Similarly, define $H_j(k+1, w, q)$ as the total expected return after the worker has migrated to the location j on the kth round, and has found the quality of location j to be q. Remembering the definitions of the set of already visited locations N^k and the set of remaining locations M^k, we can claim that the following relations [corresponding to (2) – (5) above] will hold according to the principle of dynamic programming, when $\underline{W}_{jk}$ is the optimal stopping rule for accepting a job on the kth round of onward migration.

$$V^k = \max [\, V^k(j),\, j \in M^k \,] \tag{7}$$

Again, $V^k(j)$ will be equal to

$$V^k(j) = -c_j + \int_{\underline{W}_{jk}} J_j(k, w)\, F_j(w)\, dw + V^k(j)\, [1 - \int_{\underline{W}_{jk}} F_j(w)\, dw],\ j \in M^k \tag{8}$$

For all $j \in M^k$, we will have,

$$J_j(k, w) = \int_q [qw + \beta H_j(k+1, w, q)]\, a(q)\, dq \tag{9}$$

It remains to evaluate the expected payoff from location j after k rounds of migration has occurred and the current location is j with wage w and quality revealed as q, represented as $H_j(k+1, w, q)$. This term will be the maximum of the payoffs from four options :

1) stay at the current location with total payoff $qw/(1-\beta)$,

2) local search at j with payoff $V_L(j, q)$ as given by (5) above,

3) another round of onward migration with expected total payoff V^{k+1} defined in the same way as in (7) above, and

4) return migration to one of the previous locations (an option that did not exist at the first round of migration), with expected payoff defined as z^{k+1}, evaluated below.

When an individual decides on return migration to one of the locations where he had migrated to previously, he will choose, typically, from the set of locations $i \in N^k$, $i \neq j$, when j is the current location after k rounds of migration. Of the k – 1 locations that he has visited before, he will decide to search for a job in the location which offers him the highest possible expected payoff. If he wants to search for a job in location i again, he has to incur a search cost of c_i per search as he is searching for a job in location i from outside i.e. from his current location j ; but as he has stayed at location i previously, he would know the true quality of location i to be equal to say, q_i. Thus his expected payoff from searching at location i (with possible return migration in the event of a successful search) will be the same as that of a local search in location i except the search cost is c_i, not L as in a local search, when $c_i > L$. Thus the value of z^{k+1} will be the maximum of the values $z(i, q_i)$, $i \in N^k$, $i \neq j$, when j is the current location and $z(i, q_i)$ is the expected total payoff from searching in location i for a job for possible return migration. The value of $z(i, q_i)$ will be equal to the value of a local search at location i with a search cost of c_i instead of L. Therefore we will have the following relationships [using the same ideas as in (5) and (6) above] :

For all $i \in N^k$, $i \neq j$, we have,

$$z(i, q_i) = -c_i + \int_{\underline{w}_{ci}(\underline{q}_i)} [q_i \cdot w / (1 - \beta)] f_i(w)\, dw + (1 - \int_{\underline{w}_{ci}(\underline{q}_i)} f_i(w)\, dw)\, z(i, q_i) \qquad (10)$$

When $\underline{w}_{ci}(\underline{q}_i)$ is the optimal stopping rule for search in location i from outside with search cost c_i per search and the quality level known to be q_i. Further, we have

$$z^{k+1} = \max z\,(i, q_i) \qquad i \in N^k,\ i \neq j, \tag{11}$$

and finally, the expected payoff from (k+1)th round of migration denoted by H_j (k+1, w, q) will be given by

$$H_j\,(k+1, w, q) = \max\,[qw\,/\,(1-\beta),\ V_L\,(j, q),\ V^{k+1},\ z^{k+1}] \tag{12}$$

Equations (1) – (12) describe completely my model of search and migration where return migration, onward migration and local search are three of the four options exercised by the individual in this model. Notice that the individual can migrate several times in this infinite-horizon model, he can migrate onward or backward at any time and he can also change jobs at the same location at any time. Also, the only learning that takes place in this model is about the true quality of a location after the individual migrates there. From the true value of a certain location, the individual does not try to infer what the quality level of the remaining locations might turn out to be. Such learning (which is ignored here) may occur, for instance, if the individual thinks that there is a "good" location and having found one that is not "good", becomes more convinced that the "good" location is in the set of remaining locations. Further, I have also ignored the possibility that the individual may think that the wage offer received in one location may be correlated to the unknown true quality of the location.

My main objectives in this section is to show that at every stage of the search process the individual's search/ migration strategy will follow a reservation wage property. I will do this by first showing that local search always follows the reservation wage property, as show in the following proposition.

Proposition 1

(a) For each revealed quality level q in location j the worker's local search has a reservation wage property. (b) This reservation wage for local search W_{Lj} (q) is an increasing function of the observed quality level q and a decreasing function of the local search cost L. (c) Further, the optimal value of local search V_L (j, q) and the reservation wage $W_{Lj}(q)$ can be obtained as a solution of

$$L = \int_{w_{Lj}(q)}^{\infty} \left[\{q.w/(1-\beta)\} - V_L(j,q)\right] f_{j(w)}\, dw \qquad (13)$$

$$V_L(j,q) = q.\ w_{Lj}(q)/(1-\beta) \qquad (14)$$

Corollary 1

The expected total payoff from local search V_L (j, q) is an increasing function of the observed quality q and a decreasing function of the search cost L.

Proof

Proposition 1 and Corollary 1 have been proved in Bhattacharya (1990).

Similarly, we can show that when the individual wants to search in a previously visited location i for possible return migration, his search also follows the reservation wage property.

Proposition 2

(a) For each known quality level q_i in location i, if the individual wants to follow an optimal search in location i from outside for possible return migration, his search would have a reservation wage property. (b) The reservation wage for return migration to location i is denoted by W_{ci} (q_i) and it is an increasing function of the quality level q_i and a decreasing function of the search cost at location i, c_i (c) The expected payoff from return migration to location i, denoted by z(i, q_i), is an increasing function of the observed quality q_i and a decreasing function of the search cost c_i. (d) The expected payoff z(i ,q_i) and the reservation wage W_{ci}(q_i) are obtained by solving the following two equations

$$c_i = \int_{w_{ci}(q_i)}^{\infty} \left[\{q_i.w/(1-\beta)\} - z(i,q_i)\right] f_i(w)\, dw \qquad (15)$$

$$z(i,q_i) = q_i.w_{ci}(q_i)/1-\beta \qquad (16)$$

Proof

Proposition 2 can be proved by mimicking the proofs of proposition 1 and corollary 1.

The results stated above are all standard results of search theory because both the local search and the search for return migration are standard search processes. Thus, the reservation wage in each case is a decreasing function of the respective search cost because as the search cost rises, the individual becomes more willing to accept a lower offer instead of searching more. Further, as the observed quality level rises in each case, because the additional expected benefit from another round of search is an increasing function of the quality level of a location, the individual would to accept an offer only if it pays a higher wage. Also, the equations (13) and (14) show respectively that the cost for an additional search is equal to the expected benefit from one additional search and that the expected total payoff from local search is equal to the total payoff from a job with the same quality and a wage which is exactly equal to the reservation wage. The same interpretations can be made of equations (15) and (16) in case of a search for possible return migration.

The expected total payoff from local search at location j, according to (13) and (14), is dependent on the local search cost L, and the quality level q of the location j only. Thus it is not a function of how many rounds of migration the individual has already made or how many locations remain to be visited. Similarly, when the individual considers return migration to location i, his expected payoff depends on the known quality of the location and the cost of searching for jobs from outside the location; not on how many rounds of migration has occurred in the past his current wage rate or any other aspect of his current situation. This fact enables us to analyze the individual's choice at every round of his migration process more easily, because when the individual is considering whether to migrate for (k+1)th time three of the four possible options (local search, return migration and onward migration) have expected payoffs that are independent of his current offer, and furthermore return migration and onward migration both have expected payoffs that are independent of his current wage rate and the quality of his current location. Thus although my overall search/migration model is not a standard search model, I can still show that at every round of the search/migration process the individual's optimal strategy has a reservation wage property; and at every round of migration it is possible to conclude a priori whether the individual will choose return migration over onward migration independent of his current wage rate and the quality level of his current location. These results are proved below.

Proposition 3

a) After (k-1)th round of migration, the Individual's search/migration decision for the kth round will be one of the following options with the optimal strategies for each case specified below

i) stay at the current job at wage w at the current location r whose quality level has been revealed to be q, with the total payoff $qw/ 1 - \beta$,

ii) search locally for another job with the expected payoff V_L (r, q) as given by (5) above with the reservation wage given by W_L (r, q) ;

iii) choose one of the previously visited locations $i \in N^k$, $i \neq r$, with known quality q_i, for job search for possible return migration if an acceptable offer is obtained. The optimal search strategy has a reservation wage property with the reservation wage given by W_α (q_i), and location i has the highest expected value from return migration denoted by $z^k = z(i,\ q_i)$ as given by (11), (15) and (16)

iv) choose one of the locations not visited before, say location $j \in M^k$ with unknown quality for possible onward migration to location j if an acceptable job offer is obtained. The optimal strategy is given by a reservation wage W_{kj} and the optimum expected return from onward migration is given by V^k. Further location j has the highest expected return from onward migration in the kth round given by V^k (j) and following equations determine the reservation wage and the optimum expected value of onward migration :

$$V^k = \max [V^k (i), i \in M^k] = V^k (j) \tag{17}$$

$$c_j = \int_{w_{kj}}^{\infty} [J_j(k, w) - V^k(j)]\, f_j(w)\, dw \tag{18}$$

$$\begin{aligned} V^k(j) = J_j(k, w_{kj}) = w_{kj} \int_0^{\infty} q\, a(q)\, dq\, + \\ \beta \int_0^{\infty} H_j(k+1,\ w_{kj}, q)\, a(q)\, dq \end{aligned} \tag{19}$$

When Hj (..) is given by equation (12).

Proof

After k-1 rounds of migration when the individual is in location r with a known quality level q and a job that pays w, his optimal action on the kth round will determine the expected optimal lifetime payoff starting from this round given q and w as denoted by the function H_r (k, w, q) which can be obtained from (12) above,

$$H_r (k, w, q) = \max [qw/ (1 - \beta), V_L (r, q), V^k, z^k] \qquad (12^*)$$

From (12) and from the earlier propositions, it is clear that if the individual chooses the option to stay, or search locally or return to a previously visited location, his behavior will follow the strategies specified by i), ii) and iii) of Proposition 3 above. Thus we only need to show that if he chooses the option of onward migration, then the claims made in iv) above will be satisfied.

First we show that the claims made in iv) are satisfied for the last round of migration, i.e. after the individual has migrated onward l times and has found the quality level of all locations. Assuming the last location is also called j, and q and w are the quality and wage respectively, the choice here is to stay in the current job with payoff qw/ 1 – β, or to choose local search with expected payoff V_L (j, q), or to choose return migration with expected payoff z^{l+1}; the option of onward migration does not exist - thus V^{l+1} = 0. Thus the claims made in proposition 3 obviously hold for this round.

Suppose now that the claims made in Proposition 3 hold for (k+1)th round of migration, then we will show that they hold for the kth round. From (12), we can see that as V_L (j, q), V^{k+1} and z^{k+1} do not depend on the current wage w at location j, H_j (k+1, w, q) must be a non - decreasing function of w. The optimal expected return from location j before its quality is revealed, as given by J_j (k, w) in (9) above, is related to the function H_j as follows,

$$J_j(k,w) = \int_q [qw + \beta H_j(k+1, w, q)]\, a(q)\, dq \qquad (9)$$

Therefore, J_j (k, w) is a monotonically increasing function of w. A standard result in search theory [Hey (1979)] in this case is that the optimal stopping rule $\underline{W}_{jk}$ is given by a reservation wage i.e.

$$\underline{W}_{jk} = \{w / w > w_{jk}\} \qquad (20)$$

Thus the functional equation for this round of search can be written, from (8), as

$$V^k(j) = \max_{w} \left[-c_j + \int_{w}^{\infty} J_j(k,w)\, f_j(w)\, dw + V^k(j) \int_{0}^{w} f_j(w)\, dw\right],\ j \in M^k \tag{21}$$

Of course, the reservation wage w_{kj} will maximize the right-hand side of (21). Using the fact that the right-hand side of (21) evaluated at w_{kj} will be equal to $V^k(j)$ on the left hand side, we get (18). Also, the first order condition for maximum of the right-hand side of (21) and the definition of J_j as given by (9) above will give us (19). Equation (17) holds by location j being the optimal location for return migration.

Q. E. D.

Equations (18) and (19) reveal some standard properties of reservation wages and optimal value functions that hold even for this complicated search/migration model. Equation (18) implies that the search cost for getting another offer will be equal to the increase in the expected total payoff from one additional search. On the other hand, equation (19) implies that the expected value of searching at the location j is equal to the expected return that would be obtained if an offer exactly equal to the reservation wage is accepted and an optimal search policy is followed in the subsequent rounds after the quality of location j is revealed.

Proposition 3 summarizes the optimal search/migration strategy of my model. There are certain important corollaries of this proposition that further characterize the optimal strategies. First, at the very first round, i.e. for k = 0, the optimal policy would be to migrate onward as other alternatives do not exist, and the reservation wage in that case will be w_{0i} when i is the best location for onward migration in the first stage. Thus we have :

Corollary 2

The individual's search and migration decision at the first stage (round zero), before the quality of any location is known, follows a reservation wage property with the reservation wage given by w_{0i} when i is the best location for migration in round zero Further, this reservation wage is a decreasing function of the search cost at location i and a decreasing function of local search cost L.

Proof

From Proposition 3, taking k = 0 and remembering that the only option at this round is to migrate onward we get the reservation wage w_{0i} when i is the location the individual chooses to migrate to. We also get for k = 0 and for location i the equations corresponding to (18) – (19),

$$c_i = \int_{w_{0i}}^{\infty} [J_i(0,w) - V^0(i)]\, f_i(w)\, dw \tag{22}$$

$$V^0(i) = J_i(0,w_{0i}) = w_{0i} \int_0^{\infty} q\, a\, (q)\, dq + \beta \int_0^{\infty} H_i(1, w_{0i}, q)\, a\, (q)\, dq \tag{23}$$

From (22) and (23), using the usual comparative statics methods, we can show that the reservation wage w_{0i} is a decreasing function of c_i and L. Q. E. D.

Second, the reservation wage at any round of onward migration will also be a decreasing function of search cost at that location, of local search cost and search cost at any other previously visited location, (note that "decreasing function" is used in a weak sense here i.e. non-increasing).

Corollary 3

If j is the optimal location for kth round of onward migration (k = 0, 1, 2, ...), then the reservation wage w_{kj} will be a decreasing function of c_i, i = 1, 2, 3, .. I and a decreasing function of local search cost L Further, the expected payoff from onward migration V^k (j) will also be a decreasing function of the search cost at location j, of search costs at all other locations and local search cost L

Proof

From Corollary 1, we know that the expected value of local search V_L (j, q) is a decreasing function of local search cost L and from proposition 2, the expected return from return migration z^{k+1} is a decreasing function of the outside search costs c_i, $i \in N^k$, $i \neq j$. Using the definition of H_j function as given by (12) above, H_j (k+1, w, q) then will be a decreasing function of c_i, $i \in N^k$, and of L. From (18) and (19), again using comparative statics methods, we can show that $dw_{kj}/dc_i \leq 0$, $i \in N^k$, $i \neq j$, $dw_{kj} / dL \leq 0$, dv^k (j) $/dc_i \leq 0$, $i \in N^k$, $i \neq j$ and dv^k (j) $/dL \leq 0$.

To show the effects of a change in cj, we can evaluate from (18) and (19) directly that $dw_{kj}/dc_j \leq 0$ and dV^k (j) $/dc_j \leq 0$. Finally, to show the effects of changes in the search costs of locations not previously visited, note that for the last round of onward migration, say round n, V^n is obviously a decreasing function of c_n. Thus, for the earlier round n-1, v^{n-1} can be shown to be a decreasing function of c_n. Proceeding in this way, we can show that v^{k+1} and hence H_j (k+1, w, q) is a decreasing function of c_i, $i \in M^k$. Given this, from (18) and (19), we can use comparative statics methods to show dV^k (j) $/dc_i \leq 0$ and $dw_{kj}/dc_i \leq 0$ for all $i \in M^k$. Q. E. D.

Third, since return migration, local search and onward migration are three of the four options in my model, the model is too rich to permit any stationarity property of the optimal actions of the individual. To see this clearly, note that the expected payoff from onward migration to location j at round k, V^k (j) depends on current location r through the current quality q_r because V^{k+1} and hence V^k (j) includes the expected payoff from return migration to location r in round k+1 if the individual finds the revealed quality to be too low in location j where he would have migrated onward in round k. Thus the optimal location for onward migration in round k may also be affected by the current quality q_r because the same option of returning to location r with quality qr in round k+1 may affect the payoff from migrating to location i on round k differently from the payoff from migrating to location j on round k and these return migration effects appear on all the subsequent rounds of migration too ! Therefore, V^k, the expected payoff from onward migration and the optimal location for onward migration j such that $V^k = V^k (j)$ depends on the current quality at the current location and using the same argument, the revealed qualities of past locations on the past k-1 rounds and the search costs and wage distribution of all the locations visited in the past through their potential effects on the expected payoff from return migration z^k. Can we, under these circumstances, say anything about the nature of the expected return function from onward migration because such characterizations are important for policy implications in this model ? It turns out that we can prove the following :

Corollary 4

Given a history of migration, say (for simplicity of notation) that the individual has migrated onward to location 1 on the first round (k = 0), location 2 on the second round, etc. up to location k-1 on the kth round, and given the corresponding revealed quality levels as q_1 for location 1, q_2 for location 2, etc., then we can conclude that the expected payoff from onward migration to the optimal location k-2 on round k-1 is greater than or equal to the expected payoff from onward migration to the optimal location k-1 on round k i.e ,

$$V^{k-1} (k-2) \geq V^k (k-1) \tag{24}$$

Proof

By optimality of location k-2 on round k-1, we have

$$V^{k-1} (k-2) \geq V^{k-1} (k-1) \tag{25}$$

But from the definition of V^{k-1} (k-1) and V^k (k-1) , [see equations (9) and (12) and proposition 3 above], we can see that the only difference between V^{k-1} (k-1) and V^k (k-1) is that in round k, there is the additional option of return migration to location k-2 on round k; therefore,

$$V^k (k-1) = \max \{ V^{k-1} (k-1), Z^k (k-2)\} \tag{26}$$

But we know that, as k-1 is by hypothesis the optimal location for round k, and on round k the individual had the option of searching locally for a job in location k-2 which is the current location after k-1 rounds of onward migration, thus it must be the case that the expected payoff from onward migration to location k-1 on round k was at least as great as the expected payoff from local search in location k-2 with the known quality level q_{k-2}; therefore

$$V^k (k-1) \geq V_L (k-2, q_{k-2}) \tag{27}$$

Further, as cost of local search L is less than cost of searching from outside at location k-2, c_{k-2}, the payoff from return migration must be less than the return from local search (see Propositions 1 and 2 above), i.e.,

$$V_L (k-2, q_{k-2}) > Z^k (k-2) \tag{28}$$

(26), (27) and (28) imply that $V^k (k-1) = V^{k-1} (k-1)$ and this proves (24). Q.E.D.

Corollary 4 shows that it is true that given a prospective history of migration, the expected payoff from onward migration from subsequent rounds of return migration will be decreasing. This property of the expected payoff from onward migration will be very usefull for discussing the issue of return vs. onward migration. It may also seem apparently obvious that the reservation wage for onward migration will also be a decreasing function of rounds of migration i.e. the individual will be willing to accept a smaller reservation wage for (k+1)th round of onward migration than on the kth round. This, however, is false because the reservation wage for every round depends on the optimal location for that round and the search cost and wage distribution of that location. Thus it is possible that a worker will have a higher reservation wage in a later round of onward migration.

Finally, I can use proposition 3 and corollary 4 to answer one of the most important questions that often comes up in the empirical literature on migration - if the individual chooses to migrate from a given location, when would he choose onward migration and when would he choose return migration ? On the one hand, choosing return migration over onward migration lets the individual go back to the best location visited so far with known quality; and on the other hand, choosing onward migration over return migration lets the individual explore the best location with unknown quality with the possibility that he might be able to find a high quality location.

Interestingly, the decision to choose between these two alternatives is independent of the current wage rate and currently known quality level at the current location; it depends on the sequence of quality levels that have been revealed in the locations visited earlier and of course the search costs and wage distributions of locations that have or have not been visited before. To explain this, assume that the individual has migrated onward k-1 times already, and is currently at location r with wage w and known quality q. If he decides to stay at r, or to search locally, the question of migration does not arise. But if he finds it optimal to migrate, then, from (12) and (16) above, he will migrate onward if

$$Z\ (i,\ q_i) = q_i \cdot w_{ci}\ (q_i)\ /1 - \beta < V^k \tag{29}$$

when i is the optimal location for return migration.

Therefore, as w_{ci} (q_i) is an increasing function of q_i by proposition 2 above, there exists a quality level q [i (k), k] such that (29) holds for q_i < q [i (k), k]. [Note that we put i (k) as the argument of q { i (k),k} because the optimal location for return migration i may depend on the k-1 rounds of migration that have already occurred]. As $V^k > 0$ for all k except the very last round of onward migration, q [i (k), k] > 0, thus there will always be quality levels of previously visited locations that are so low such that the individual will always choose onward migration over return migration if he decides to migrate at all. For q_i > q [i (k), k] the individual will always choose return migration over onward migration if he decides to migrate at all.

Further, because of corollary 4, given any prospective history of migration, as more and more rounds of onward migration occur, the expected payoff from subsequent rounds will decrease; while by definition of return migration the expected payoff from return migration will increase as more and more rounds of migration occur. Thus, either the individual will never choose return migration at all, or there will exist round k* such that he will choose onward migration up to round k*–1 and on the k*th round he will choose return migration for the first time. As the choice of return migration coincides with the end of the search/ migration process, k* is the last round of migration that will be observed. The last round k*, of course, depends on the entire past history of onward migration and the associated revealed quality levels and the actual wage offers in the past rounds, so we can not a priori determine the last round of onward migration by only looking at the search costs and wage offer distributions.

Thus we have proved an important consequence of the optimal search/migration decision of the individual to determine whether he will choose return or onward migration :

Proposition 4

Given a prospective history of onward migration for k-1 rounds, if staying at the same job or searching locally for another job is not optimal, the individual will decide to migrate either by choosing return migration or by choosing onward migration If location l (with known quality level q_l) is the best location for return migration on round k, then there will exist numbers q [l (k), k] > 0 such that

i) If q_l > q[l (k), k], the individual will choose return migration over onward migration on round k,

ii) if q_l < q [l (k), k] the individual will choose onward migration over return migration on round k,

ii) q [l (k), k] is determined by z (l, q_l) = v^* [see (29) above].

Further, given a prospective history of migration and associated revealed q and wage offers, there will always exist a critical round k* such that for k < k*, the individual is observed to choose onward migration and for the k*th round the individual will choose return migration thereby ending the search/ migration process. The critical round k* depends on the migration profile, at k* the expected payoff from return migration exceeds that from onward migration for the first time

Proposition 4 implies that the individual will always choose onward migration so long as the quality of the locations already visited is low enough so that it is not justifiable to choose return migration; further the choice of return vs. onward migration does not depend on the revealed quality of the current location or the wage in the current location. The current revealed quality and the current wage determine, along with other parameters of the model, whether the individual will choose to migrate at all, whereas the expected payoff from return migration compared with that of return migration determines which type of migration will be chosen.

3. Properties of optimal search/migration decisions

In this section, I will show under what conditions the individual will choose the different options available to him. As shown above, onward migration will be chosen for all rounds in my model except the last round k^*, $(0 < k^* < I-1)$, i.e. the individual will either never choose return migration, or choose return migration once and terminate the search/migration process. Further, at each round, if he chooses local search or if he decides to stay in the current job, he also ends the search/ migration process.

Assume that the individual has already migrated onward for k-1 rounds $(k = 0, 1, ..)$, he is currently now in location r, with current wage w_r and known quality q_r. Further, assume that his best prospective location for return migration happens to be location i with known quality level q_i with expected payoff $z(i, q_i)$, and also the best prospective location for onward migration happens to be j on round k, with expected payoff $V^k(j)$.

On round k, given w_r and q_r, his choice among the four options depend on which option determines the value of $H_r(k, w_r, q_r)$.Recall that the optimal expected payoff from the options available in round k is given by (12*) above which can be written as

$$H_r(k, w, q) = \max[qw/(1-\beta), q_l w_{Lr}(q_l), V^k(j), z(i, q_i)] \qquad (30)$$

As we saw in proposition 4 above, the individual will choose onward migration if q_i is less than a given number $q[i(k), k]$. For the first part of this section, then, I will explain the nature of the optimal policies assuming that onward migration is preferred to return migration as $q_i < q[i(k), k]$. Note that this does not indicate that onward migration will be chosen by the individual under all circumstances. On the contrary, it merely implies that if the individual chooses any kind of migration over local search and no search options, he will always choose onward migration and not return migration.

To characterize the individual's optimal decisions when *onward migration is the favoured over return migration*, define q^* by

$$\frac{q^{*}\ .w_{Lr}(q^{*})}{1-\beta} = V^{k}(j) \qquad (31)$$

Thus q* is that value of revealed quality in location r such that if the worker decides to search locally with revealed quality q_r, his expected lifetime income from local search is exactly equal to his expected lifetime income from migrating to the best location j on round k.

As w_{Lr} (q) is an increasing function of q, the curves

$$q_r \,.\, w_r / (1-\beta) = V^k (j) \qquad (32)$$

and

$$w_r = w_{Lr} (q_r) \qquad (33)$$

intersect at $q_r = q^*$, $w_r = w_{Lr}$ (q*) because of (31). Figure 1 shows how the search and migration decisions are made on round k when onward migration is preferred to return migration. The negatively sloped curve labeled V^k (j) on Figure 1 is the rectangular hyperbola given by (32) above, which shows at what current wage and quality levels staying in the current job in location r has the same expected payoff as onward migration to location j. The positively sloped line graphs equation (33) and shows the combinations of w_r and q_r such that staying at the same job in location r will have the same expected payoff as local search at location r. The other negatively sloped line marked z (i, q_i) graphs the equation

$$w_r . q_r / (1-\beta) = z (i, q_i) \qquad (34)$$

i.e. the combinations of current wage and quality levels that cause staying at the current job at location r to have the same expected payoff as returning to the best location i. As onward migration is preferable to return migration as we assumed q_i < q [(i) (k), k] for this part of this section, the curve marked z (i, q_i) is below the other curve V^k (j). On Figure 1, the optimal decisions of the individual can be easily analyzed if we notice that to maximize his total payoff which is in the multiplicative form w.q / 1−β, he needs to choose an option that will take him as far northeast on the diagram to the highest rectangular hyperbola of the form $q_r.w_r$ /1–β. Thus, suppose on the k-1 round he has migrated to location r with a wage w_r and finds the quality of location r to be q_r such that he finds himself at point c in Figure 1. If he decides to stay in the current job, his lifetime payoff is $q_r.w_r$ / (1–β); if he opts for local

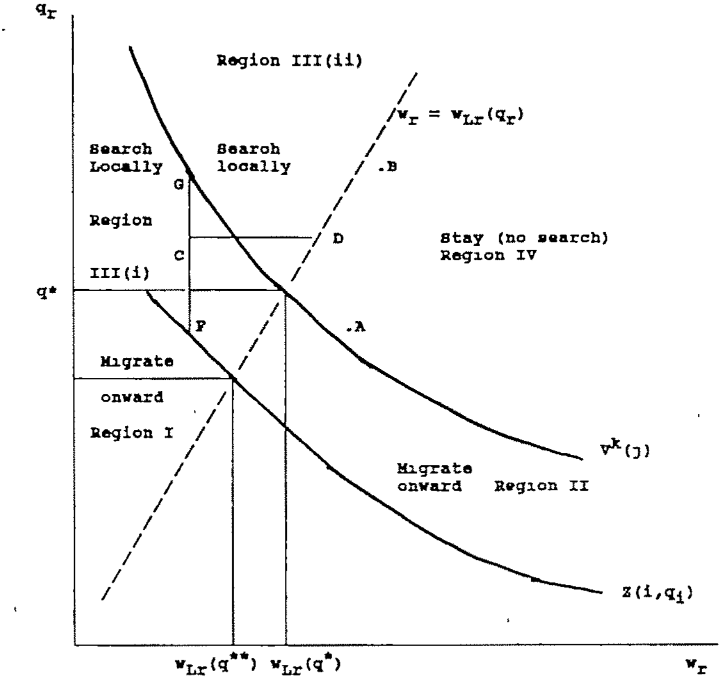


Figure 1

Onward migration is preferred to return migration

search he can retain the quality level q_r and move horizontally to the positively sloped line with an expected payoff given by w_{Lr} (q_r) .q_r / (1 – β), i.e. to point D on Figure 1. If he chooses return migration, then his payoff is given by z (i, q_i) which is shown by the curve marked z(i, q_i); and finally, if he chooses onward migration, his expected payoff is V^k (j) shown by the corresponding curve on Figure 1. It is obvious that if c is the current position after k–1 rounds of onward migration, then the optimal policy is to stay at location r, and search locally for a job with a reservation wage of w_{Lr} (q_r), thus ending the search /migration process at this round. In general, from Figure 1, we can identify the following regions (recall that we have assumed onward migration to be preferred to return migration for this part).

Figure 1 is divided into four regions.

Region I : (Low wage, low observed quality).

If $q_r < q^*$ and $w_r < w_{Lr}$ (q^*), then the optimal policy is to search for a job at location j and migrate onward to that location.

Region II : (very low level of observed quality relative to the current wage level).

If $w_r > w_{Lr}$ (q^*) but q_r is low enough such that $[q_r.w_r / (1 - \beta)] < V^k$ (j), the optimal policy again is to search in location j and migrate onward. Here, even though the current wage rate could possibly be high, the observed value of q_r is so low that it is worthwhile to find a job in another location whose quality is unknown.

Regions III (i) and III (ii) (High revealed quality, relatively low wage

If $q_r > q^*$ but $w_r < w_{Lr}$ (q^*), the optimal policy is to search locally for a better job. Note that w_{Lr} (q_r) is an increasing function of q_r, i.e. given a higher value of observed quality the worker will decide to search locally and quit his present job even though the current job pays a higher wage rate. This is because, with an assured level of high quality, even a moderate chance of finding a better job is worth the local search cost.

Region IV (High wage and moderate quality level)

If $w_r \geq w_{Lr}$ (q^*) and $q_r w_r / (1 - \beta) > V^k$ (j), the individual decides not to search at all either locally or outside location r. The decision to stay may arise even at very levels of quality (point A on Figure 1) if the wage rate is high enough. From a point like A, if the observed quality level becomes a little lower, the worker will try to migrate to another location in spite of the high wage rate. On the other hand, the worker might decide to stay in the current job at a point like point B on Figure 1 where wage is high and quality level is moderate. Were the quality level a little higher, he would have searched locally for a better job. Although this conclusion may appear to be surprising, it is expected in this model because a higher assured quality level makes the search for an even better job (paying a higher wage than the currently well paying job) justified. Thus a higher quality level will make local search more attractive even if the current job pays well.

Coming now to the other possibility where return migration is preferred to onward migration, define q** as

$$q^{**} . w_{Lr}(q^{**}) / 1 - \beta = z(i, q_i) \qquad (35)$$

On Figure 2, we have a diagram identical to Figure 1, except return migration is preferred to onward migration as the curve marked $z(i, q_i)$ is to the right of the curve marked $V^k(j)$ and the critical quality and wage levels are q^{**} and $w_{Lr}(q^{**})$ respectively. Again, the regions are, as indicated on Figure 2 : Region I - return migration is optimal - both current quality and wage are too low.

Region II - return migration is optimal - current quality is too low relative to current wage.

Region III (i) and III (ii) - local search is optimal.

Region IV - staying at the same location at the same job is optimal.

As before, the same comments made in the analysis of onward migration are also valid here. This model, therefore, has some very interesting implications for job quits and migration behavior. First, for a relatively low quality location ($q < q^*$, or $q < q^{**}$ as appropriate), an increase in wage will discourage both migration to other locations and local job turnovers, but an increase in location quality may increase turnover in the local job market because more workers will start searching locally. Second, for jobs with moderate wages, the quit rate will increase sharply if the observed location quality rises or falls by relatively small amounts. If the quality level rises, the workers quit and search locally and if the quality level falls, they will migrate.

Third, suppose the (k-1)th round reservation wage is $w_{(k-1)}r$

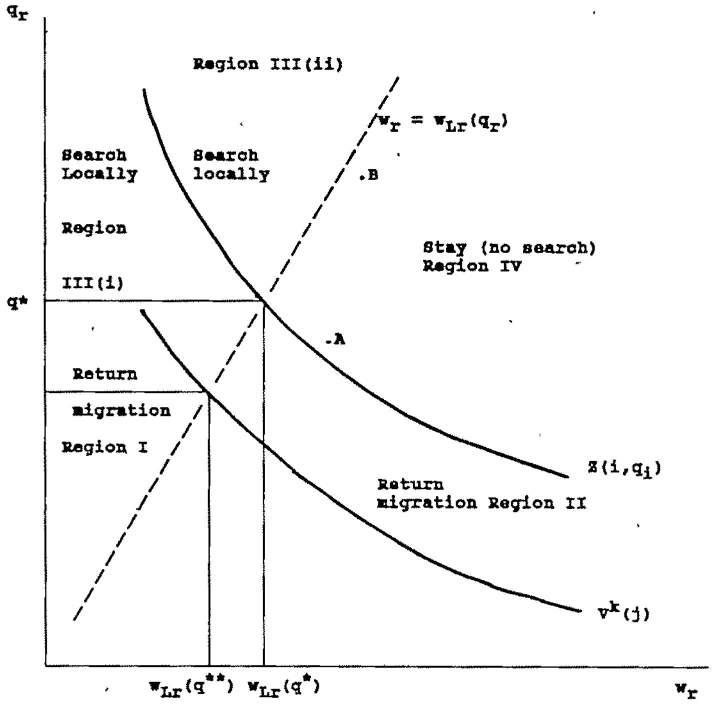


Figure 2
Return migration is preferred to onward migration

(see proposition 3 above). This can possibly be below the critical wage rate w_{Lr} (q^*) or, if return migration is preferred, below w_{Lr} (q^{**}). Note that the individual knows, while deciding to move to location r in round k-1, whether he will return or move onward in round k after the quality level of location r is revealed (and provided that he does not want to stay in r or search locally in r). In this case, we may observe workers accepting jobs where they know that after migrating to location r on round k-1 and accepting a job that pays w_r [between $w_{(k-1)r}$ and w_{Lr} (q^*)], they will certainly quit after working for one period. Indeed, these jobs can be characterized as "one-period-jobs" and the individual accepts these jobs in order to migrate to a particular location. After working at such a job for one period, if the revealed quality of the location turns out to be low (less than q*), he quits to migrate onward and if the quality level turns out to be high (greater than q*), he quits to search locally.. Note that if a location has a large number of these jobs, any economic policy that improves location quality will not change the quit rate substantially; it will only reduce the rate of outmigration and increase the rate of local job turnovers.

3. Search costs, job opportunities and migration decisions

I have been able to find the effects of two kinds of parameter changes in this model. Recall that there are two types of search costs here. The first type is c_i, i = 1, ..., I, which is the cost of searching for a job in location i from outside location i. By assumption, this search cost is applicable when an individual is searching for a job in location i from outside for the first time or with the purpose of return migration. The second type of search cost is L, which is the search cost for local search, assumed to be the same for all locations. I have shown the effects of changes in both types of search costs in the following propositions. Assume, as in the last section, that the current location is r after k-1 rounds of onward migration, the current wage and known quality levels are w_r and q_r respectively the best locations for prospective onward and return migration are j and i respectively, and onward migration is currently preferred to return migration. Proofs of all propositions in this section are in the Appendix.

Proposition 5 (a)

An Increase in the cost of search c_j for location j, which is the optimal location for round k, with all other search costs constant, will reduce the reservation wage for the earlier round k-1 where the individual decides to migrate onward to location r (b) It will also reduce the region of wage offers which the Individual considers as a region of one-period jobs. (c) Further, even with lower revealed quality, for some jobs with low wages the worker will be encouraged to search locally in location r instead of migrating to location j (d) Similarly, for some jobs with high wages but lower revealed quality, the Individual will stay at the current job at location r, rather than choosing to migrate to location j. (d) Finally, if the rise in this search cost is large enough, there will be a switch from onward migration to return migration

If there is an increase in local search cost L in all locations, then the expected payoff from onward migration at every round V^k falls (see corollary 3 above). Also, for each observed quality, reservation wage for local job search w_{Lr} (q) and hence the expected payoff from local search V_L (r, q_r) will also decrease. Therefore the net effect on the reservation wage w_{kj} is ambiguous. However, if we assume that an increase in local search cost affects the expected value of local search V_L (r, q_r) for all q_r in location r more than the expected value of the entire search and migration program for the next round of onward migration to location j i.e. V^k (j), then we can conclude.

Proposition 6

An increase in local search cost will reduce the range of "one period-jobs" by reducing the value w_{Lr} (q^*) For some high quality low wage jobs the individual will be induced to migrate again rather than search locally, and for some high quality high wage jobs the worker will be induced to stay at the current job rather than search locally at the same location. The reservation wage on round k-1 will fall as a result of a rise in the local search cost

We can also find the effects of an improvement of job opportunities in location on the search/migration decisions of the individual as follows. Let F_j (w), the wage distribution in location j, be parameterized as F_j (. |α_j). Then we can show the effect of an improvement in job opportunities in location j on the search and migration decisions in round k-1 when r is the location for round k-1.

Proposition 7

Let F_j (. |α) stochastically dominate F_j (|α') whenever $\alpha > \alpha'$. Then the reservation wage for round k-1, $W_{(k-1)\,r'}$ will be higher for $\alpha_j = \alpha$ than the corresponding reservation wage when $\alpha_j = \alpha'$ Further, for some high quality low wage jobs [region III (I)], the individual will decide to search locally at a higher quality for $\alpha_j = \alpha$. For some low quality high wage jobs, (region II), the individual will decide to stay at the current job (instead of migrating again) at a higher quality when $\alpha_j = \alpha$ Finally, the range of "one period jobs" will be greater for $\alpha_j = \alpha$.

Finally, it is possible that for $\alpha_j = \alpha$, onward migration is the preferred option, while for $\alpha_j = \alpha'$ return migration is preferred to onward migration

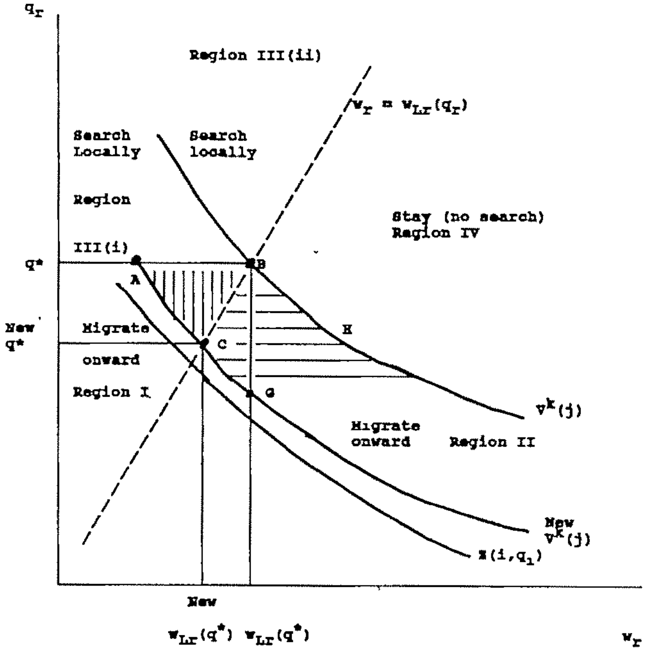


Figure 3

Effects of change in search cost – onward migration continues to be preferred to return migration

Proposition 7 considers the effects of an improvement in job opportunities at the location j where the individual intends to migrate on round k, on the migration decisions in the last location r. I now consider the effect of an improvement in job opportunities at the current location r. If $F_r(\,.\,|\alpha)$ stochastically dominates $F_r(.\ |\alpha')$, the effects will be as follows

Proposition 8 an improvement in the job opportunities in the current location with job opportunities at the other locations remaining unchanged, will increase the range of one-period jobs by increasing $w_{Lr}(q^*)$, will cause a shift from onward migration to local search for some low quality low wage jobs, and a shift from no search to local search for some high quality high wage jobs. However, if onward (return) migration was the preferred option before the improvement of job opportunities in location r, it will continue to be the preferred option even after the improvement. It will also raise the reservation wage in location r on round (k-1).

5. Concluding remarks

I have developed a fairly elaborate model in this paper to analyze the individual's options in his dynamic migration / search process, with four options – local search, onward migration, return migration and no search – considered by the individual at every round of the infinite – horizon model. There are a large number of conclusions that emerge from this model that will possibly be helpful in understanding the issues of onward and return migration, job quit behaviors, local job search, and how "quality of life" in a region can play a role in shaping the flow of migration in and out of that region. Furthermore, most of these conclusions will have testable implications that may be important for empirical work.

a) Return and onward migration – This model shows that the individual will typically follow a pattern of successive rounds of onward migration before choosing return migration to the "best" previously visited location and ending the migration process. Both onward and return migration decisions follow a reservation wage property, but the reservation wage for return migration to location i w_{ci} (q_i) depends only on the quality and the search cost parameters of location i only, and it is independent of any characteristics of the current location. However, the reservation wage for onward migration w_{kj} is a function of the known quality and the current wage (and other parameters) of the current location r, as well as a function of the relevant parameters of the destination location j. Thus if a region experiences a lot of outmigration most of which consists of workers returning to their previous locations, then a small change in current quality or current job attributes will hardly change the flow of migration from that region. Only a substantial improvement in quality or job opportunities will make the workers choose either the local search or no search option and change the flow effectively. On the other hand, if a region has mostly onward migration as the main part of outmigration, then a small change in either q or w will change the workers' reservation wage and will have some impact on outmigration.

b) Local job search and job quit behavior - most of the standard search - theoretic literature on the job search and job quits [see for example, Burdett (1978), Wilde (1979), Mortensen (1984) etc.] consider the worker's job search behavior based on disappointment with the current job or willingness to explore other career opportunities as the primary force behind on-the-job search. However, as this paper shows, there are two distinct parts of the on-the-job search process – search for jobs in the same location and search for jobs in a new location leading to eventual migration. As shown in sections 3 and 4 above, for high quality locations, even with relatively high wage jobs local search is often the preferred option, while for low quality locations, onward or return migration is the preferred choice unless the wage level is very high. Thus in high quality locations, quit rates are going to be dependent only on location - specific search and wage distribution parameters and thus more responsive to regional policies and hopefully more stable over time. On the other hand, in low quality locations, quit rates are going to depend on parameters of the current locations and on the job opportunities and search costs of other locations as well - thus they will probably be more unstable over time and less responsive to local policy changes.

Further, this paper brings into focus the existence of one period jobs which the individual is certain to quit after working for one period only. Whereas in high quality areas such one period jobs are instrumental in bringing in new workers that are probably going to look for jobs locally after quitting, for low quality locations a large number of these jobs can lead to serious instability in the labor force because of the presence of a large group or transitional workers. In this case, surprisingly, if correct information about the poor quality of the location is available to potential migrants, this will reduce the quit rate and the instability of the labor force substantially, whereas improving quality of life by a small amount may not be helpful in reducing the quit rate.

The conclusions obtained in sections 3 and 4 show that strong testable hypotheses may be derived from my model about the relationships between duration of migration, return and onward migration, local and national reservation wages and local search and remigration decisions. One can also use the model for analyzing how job turnovers, job quits, and migration rates are affected by changes in local or national search costs and job opportunities at different locations. Recent empirical work on migration [Herzog and Schlottman (1983), Schlottman and Herzog (1982), Morrison and Davanzo (1987), Davanzo (1983), Allen (1979) etc.] have shown that return migration and onward migration are two different parts of the aggregate migration flow in any given year and the worker's personal characteristics, disappointment with the current location and the nature of information about other job opportunities are important determinants of which type of migration will be adopted. Recent empirical work on search models, job quits,labor turnover, etc. also demonstrate that attractiveness of the current location and / or the current job is an important part of the job search process. I have developed a model where both these issues can be analyzed in terms of a single search and migration model with uncertainty about locational characteristics. The conclusions in section 4 can be very useful to determine under what conditions the individual will adopt onward or return migration or local search and what parameter changes will lead him to switch between these actions. Similarly, the results of this paper can also be used to examine the effects of changes in local and national economic policies affecting job opportunities, job information, locational characteristics, etc. on the workers' return and onward migration decisions.

This paper introduces two important technical aspects to the large number of papers in the field of sequential decision - making under incomplete information (there are two types of papers published in this area - the papers on job search with incomplete information [Berninghaus (1986), Lippman and McCall (1981), McCall and McCall (1981), Weitzman (1979), Wilde (1979), Mortensen (1984) etc.]; and papers on the analysis of purchase of experience goods [Wilde (1980), Hey and Mckenna (1983), etc.]. First, one new feature my model is the searcher's option of searching on a conditional distribution with lower search cost but with complete information or searching on the entire distribution with incomplete information - the option of local search. Therefore, belated information, [see Lippman and McCall (1981)], even if favorable, can cause the worker to quit his job and to search locally in my model, while this option does not exist in the papers cited above. Second, in my model the individual has the option of return migration which is the same as searching in a previously visited location i.e. on a conditional distribution with complete information but with higher search cost. This feature, again, is absent from the papers cited above and this modified recall option causes the optimal policies to be dependent on the entire history of past migration and the qualities of locations observed in the past. Of course, this paper uses a model with a discrete number of locations, whereas most of the above papers analyze search with a continuous two-dimensional distribution where only one random variable is observable in the first stage of search.

McCall and McCall (1987) considers the same problem of sequential search/migration decisions in multi-armed bandit framework. Roughly speaking, the paper considers the following problem (in the notation of my paper) – the individual faces a choice of migrating to one of the I locations. In a typical location i, there are a *finite* number of wage offers w_{i1}, w_{i2}, ... w_{ig}. The individual spends a search cost c_i to search for a job in a location. If an offer is acceptable, he migrates to i, works for one period, earns w_i and finds the quality of the location to be either of the *two* possible values α_i or $-\alpha_i$ (each with probability 1/2). At this stage, he has *three* options - stay at the current job, migrate onward, or return to a previously visited location. If he decides to return to location i where he had a job with wage w_i before and where the quality was (say) α_i, he does not need to search for a job in location i again; he gets another job with the same wage w_i. However, if he decides to go back to i, he only knows the value of α_i, thus quality of location i may still turn out to be α_i or $-\alpha_i$, even though he had already visited i before and observed the quality to be α_i. This model also has a moving cost k_i to be spent every time a migration occurs. It is obvious that my model is possibly more general than theirs, because although I do not consider an explicit moving cost, I consider an *infinite number of values* of both wage and location quality. More importantly, however, I consider the option of local search which is ignored in their paper. Further, in case of return migration in my model, the individual knows the quality of the previously visited location with certainty but he has to search again for a job at a previously visited location – this is just the opposite of their paper where they have a given wage and unknown quality in case of return migration. Although it is an empirical issue to determine what happens in case of return migration, it seems more reasonable to assume the quality of the previously visited location to be given, and the wage uncertain because the worker only knows the wage distribution in that location and he has to search for a job there again. Finally, in spite of the fact that in several ways my paper is more general and possibly more reasonable, there is a large number of testable conclusions obtained in my paper (see section 3 and 4) whereas McCall and McCall (1987) obtains very few of these results.

I conclude this paper with the following remarks –

a) The model can be easily modified for workers who currently hold jobs in locations with known location quality, simply by starting the analysis from the second round i.e. $k = 1$. All the conclusions obtained in sections 3 and 4 will continue to be valid in this case.

b) The basic results of the model, namely, existence of reservation wage in different situations and the characterizations of search/migration policies shown in figures 1 and 2, can be retained even if we assume that quality of locations is differently distributed across locations i.e. A (q), the probability distribution of quality, is possibly different for different locations. Further, these results will also continue to hold if we relax the assumption of independence of the wage and quality of locations, or assume an additive (instead of a multiplicative) form of the payoff function or assume different types of local search costs for different locations. As expected, if any of these assumptions are relaxed, it becomes difficult to derive the additional properties of the search / migration policies shown in section 4.

c) Finally, learning about the quality of the current location or obtaining a job in a location may signal to the individual some properties of quality or wage distribution of the same location or of another as yet unvisited location. Such learning may alter the sequence of migration decisions that have been shown to be optimal in this paper. If such learning is significant in an individual's migration decision, the prospective employers can use some sort of signalling mechanism to attract individuals with the characteristics that they prefer. Equilibrium in such a market will then determine, along with the flows of migration, the wage differences that will exist between high and low quality locations. Problems like this, which are ignored here, will be analyzed in another paper in the future.

Appendix

Only propositions 5 and 8 will be proved in this appendix. Proofs of other propositions can be constructed on similar lines.

Proof of Proposition 5

From corollary 3 we know that $dV^k(j) / dc_j < 0$.

Coming back to location r, note that

$$H_r(k, w, q) = \max [qw / (1 - \beta), V_L(r, q), V^k, Z^k] \qquad (12^*)$$

so that as c_j goes up and $V^k(j)$ falls, q* as defined by (33)
as

$$\frac{q^*.w_{Lr}(q^*)}{1-\beta} = v^k(j) \qquad (31)$$

will fall as c_j goes up and $w_{Lr}(q^*)$ will also fall.

Figure 3 shows that as there is a downward shift in $V^k(j)$, the region of "one-period-jobs" will become smaller as $w_{Lr}(q^*)$ falls. This proves part (b) of proposition 5. For jobs in the area ABC on Figure 3, the (k-1)th round decision will shift from onward migration to local search. This proves part (c) of Proposition 5. Similarly, for jobs in the area BCHG, the individual will shift from migration to no search in the second stage. This proves part (d). As c_j rises $H_r(k, w_r, q_r)$ will stay constant for some w_r and q_r, and will fall for others. From (18) and (19), substituting r for j and k-1 for k we get the optimality conditions as

$$C_r = \int_{w(k-1)r}^{\infty} \left[J_r (k-1, w) - V^{(k-1)}(r) \right] f_r(w)\, dw \qquad (36)$$

$$V^{k-1}(r) = J_r\left[(k-1), w_{(k-1)r}\right] = w_{(k-1)r} \int_0^{\infty} q\, a(q)\, dq +$$

$$\beta \int_0^{\infty} H_r\left[(k), w_{(k-1)r}, q\right] a(q)\, dq \qquad (37)$$

Therefore, if we assume (0, ∞) being the support for the probability distribution of q, we can show from (35) and (36) that

$$\frac{dJ_r(k-1, w_r)}{dc_j} < 0 \qquad (37)$$

and

$$\int_{w_{(k-1)r}}^{\infty} \frac{dJ_r(k-1, w_r)}{dc_j} f_r(w_r)\, dw_r < 0 \qquad (38)$$

Using comparative statics methods in (18) and (19), we can show now that

$$\frac{dw_{(k-1)r}}{dc_j} < 0, \qquad \frac{dV^{k-1}(r)}{dc_j} < 0 \qquad (39)$$

This proves parts (a). Finally, it is easy to see that if the rise in c_j is large enough, the downward shift in V^k (j) may be large enough to send it below the z (i, q_i) curve in Figure 3; in this case return migration would be preferred to onward migration.

Q. E. D

Proof of Proposition 8

I will treat α as a continuous variable, assuming that $F_r(.|\alpha)$ stochastically dominates $F_r(.|\alpha')$ whenever $\alpha > \alpha'$. From (13) and (14), by using partial integration we can write (substituting r for j),

$$L = \int_{w_{Lr}(q)}^{\infty} \{[q.w/(1-\beta)] - V_L(r, q)\} f_r(w)\, dw \tag{40}$$

$$V_L(r, q) = q.\ w_{Lr}(q)/(1-\beta) \tag{41}$$

By using comparative statics methods and noting that $\delta F_r(w \mid \alpha) / \delta \alpha < 0$ because of stochastic dominance, we can show

$$\frac{dV_L(r, q)}{d\alpha} > 0, \qquad \frac{dw_{Ln}(q)}{d\alpha} > 0 \tag{42}$$

It is now routine to show from (12*) that $\delta H_r(k\text{-}1, w_r, q_r) / \delta\alpha > 0$ and $\delta J_r(k\text{-}1, w_r) / \delta\alpha > 0$. From (18) and (19), we can again show that $\delta w_{(k-1)\,r} / \delta\alpha > 0$ or the reservation wage for the earlier round rises as α rises. The rest of the results can be obtained by examining (18) and (19) and observing that in Figure 1, there is a rightward shift in the curve $w_r = w_{Lr}(q_r)$.

Q. E. D.

REFERENCES

Allen, J. (1979), "Information and Subsequent Migration: Further Analysis and Additional Evidence" *Southern Economic Journal*, April 1274 – 84

Berninghaus, S. (1986), "Job Search with Belated Information and wage signalling," *Journal of Economic Dynamics and Control*, 495 –

Bhattacharya, G (1990), Migration Under Uncertainty about Quality of Locations", *Journal of Economic Dynamics and Control* 14 (1990), 721 – 739

Burdett, K (1978), "A Theory of Employee Job Search and Quit Rates", *American Economic Review*, 68, 212 – 20.

DaVanzo, J. (1983), "Repeat Migration in the United States · Who Moves Back and Who Moves on," *Review of Economics and Statistics* 552 – 559.

Herzog H. W and A Schlottmann (1983), "Migrant Information, Job Search and the Remigration Decision," *Southern Journal of Economics*, July, 43 – 56

Herzog, H W , R A Hofler and A Schlottmann (1985), "Life on the Frontier : Migrant Information, Earnings and Past mobility", *Review of Economics and Statistics* 67, 373 – 82

Hey, J D. and C J McKenna (1981), "Consumer Search with Uncertain Product Quality," *Journal of Political Economy*, 89 (1), 54 – 66

Katz E. and O Stark (1984) "On Migration in the Presence of Asymmetric Information" *American Economic Review* 74, 533 – 534.

Katz, E and O Stark (1987), "Migration, Information and the Costs and Benefits of Signalling", *Regional Science and Urban Economics* 17, 323 – 331.

King, R. (1986), Return Migration and Regional Economic Problems Croom Helm London.

Lippman, S and J. McCall (1981), "The Economics of Belated Information," *International Economic Review*, 22, 135 – 146

McCall, B P and John McCall (1981), "Systematic Job Search, belated Information, and the Gittins Index," *Economics Letters* 8, 327 – 333

McCall, B P and John McCall (1987), "A Sequential Study of Migration and Job Search", *Journal of Labor Economics* 5, 452 – 476.

Morrison, P. A and J DaVanzo (1987), "The Prism of Migration: Dissimilarities Between Return and Onward Movers," *Social Science* Quarterly, 504 – 515.

Mortensen, D. T., (1984), "Job Search and Labor Market Analysis,' Discussion Paper no 594, The Centre for Mathematical Studies in Economics and Management Science, Northwestern University

Schlottman, A M and H. W Herzog (1981) "Employment Status and the Decision to Migrate,' *Review of Economics and Statistics*, November, 590 – 98

Schlottman, A. M. and H. W. Herzog (1982) "Home Economic Conditions and the Decision to Migrate : New Evidence for the U. S Labor Force," *Southern Economic Journal*, April, 950 – 61.

Weitzman, M. (1979), "Optimal Search for the Best Alternative,' *Econometrica* 47, 641–654.

Wilde, L. (1979), "An information-theoretic approach to Job Quits." In Studies in the Economics of Search, edited by S. A Lippman and J. J McCall, Amsterdam, North Holland

Wilde, L. (1980), "On the Formal Theory of Inspection and Evaluation In Product Markets," *Econometrica*, 48, 1265 – 1279.

Monetary Uncertainty, Velocity and the Nominal Interest Rate : An Expository Note

Parantap Basu[1]

1. Introduction :

A basic empirical fact about the behavior of the income velocity of circulation in the U. S. is its dramatic decline during the 1980s following the change in the operating procedures of the Federal Reserve. This unprecedented and unexpected decline in velocity was interpreted by the Keynesian school as demise of monetarism. Monetarists [e.g. Friedman (1984)] object to such inferences, arguing that the sharp decline in velocity in 1982 was the result of increased monetary uncertainty caused by a change in the operating procedure of the Federal Reserve over the 1979 to 1982 period.

The object of this paper is to re-examine Friedman's hypothesis about the role of uncertainty in monetary growth in causing decline in income velocity in a general equilibrium framework. Using a cash-in-advance model a la Lucas and Stokey (1987), I characterize the effects of an increase in uncertainty in the monetary growth rate on the income velocity and the nominal interest rate. Faced with a choice between "cash goods", which are cash-in-advance constrained, and "credit goods", for which invoices are issued by the seller, the typical consumer substitutes "cash goods" for "credit goods" in the presence of increased uncertainty in monetary growth. This happens because an increase in monetary uncertainty raises the real return on money which acts as a tax on the holding of credit goods. As a result, in equilibrium the income velocity declines and the nominal interest rate, which is inversely related to the proportion of "credit" to "cash" goods, also unambiguously falls. The extent of the decline in velocity depends crucially on the elasticity of substitution between cash and credit goods. Overall, this result is consistent with Friedman's monetary uncertainty hypothesis about the decline in velocity.

---

[1] I would like to acknowledge Rick Geddes for his comments on the first draft of this paper.

2. The Model

I consider a parametric version of a cash-in-advance model as in Lucas and Stokey (1987). In the Lucas and Stokey model, the trade in a full range of Arrow-Debreu securities is prevented by a friction which is represented by a category of goods (C_{1t}) that requires cash payments. The other category of goods (C_{2t}) allows trade on the basis of credit payable in cash by the consumer at the end of the trading day. At the beginning of date t, the representative agent has a choice between the use of "cash goods" (C_{1t}) and "credit goods" (C_{2t}) which are the only two available goods selling at the same price p_t. There is no capital accumulation and the technology in each period is simply $C_{1t} + C_{2t} \leq Y_t$. Money is held because without it cash goods at any date cannot be purchased. At the beginning of each date (say, t) agents first receive a monetary transfer, $\overline{M}_t - \overline{M}_{t-1}$ and interest income from the redemption of previously held securities. They then conclude their trade in securities. The goods market then opens where the agent uses the balance to purchase "cash goods." Any left over cash and the revenue from the sale of endowments finance the purchase of "credit goods". The remaining cash balance is then carried forward to the next period. The only uncertainty in the economy arises from stochastic money supply growth. At each date, the monetary growth rate is revealed before the security trading takes place.

The representative consumer, therefore, solves the following optimization problem :

$$\text{Max } E_0 \sum_{t=0}^{\infty} \beta^t U(C_{1t}, C_{2t}) \tag{H}$$

$$\text{s.t. } M_t^1 + q_t B_t \leq M_{t-1} + B_{t-1} + \overline{M}_t - \overline{M}_{t-1} \tag{1}$$

$$P_t C_{1t} \leq M_t^1 \tag{2}$$

$$P_t C_{2t} + M_t \leq (M_t^1 - P_t C_{1t}) + P_t Y_t \tag{3}$$

where M_{t-1} = cash holding at the beginning of date t,

q_t = price of a nominal bond which pay \$1 next period,

B_t = quantity of nominal bonds purchased at date t,

M_t^1 = cash holding at date t after bond trading is concluded,

M_t = cash balance at the end of period t,

β = constant discount factor $\in (0, 1)$,

$\overline{M}_t$ = money supply at date t.

Assumption 1 :

The utility function $U(C_1, C_2)$ is of the following form:

(a)
$$U(C_1,C_2) = [\phi(C_1,C_2)]^{1-\gamma}/(1-\gamma) \tag{4}$$

where
$$\phi(C_1,C_2) = [\alpha C_1^{-\rho} + (1-\alpha)C_2^{-\rho}]^{-1/\rho} \tag{5}$$

with $0 < \alpha < 1$ and $0 \leq \rho < \infty$

Assumption 2 :

The money supply, $\overline{M}_t$ follows the process:

$$\overline{M}_{t+1} = \omega_{t+1}\overline{M}_t \tag{6}$$

where $\{\omega_t\}$ is a serially uncorrelated stochastic process.

Assumption 3 : The endowment y_t grows at a constant rate α.

Assumption 1 is the main feature of the parametric form used in this paper. Note that γ can be treated as the risk aversion parameter, while ρ is inversely related to the elasticity of substitution between cash and credit goods. Assumptions 2 and 3 are made for the sake of tractability.

The first order conditions for this problem are given by :

$$U_2(C_{1t},C_{2t})/P_t = \beta E_t[U_1(C_{1t+1},C_{2t+1})/P_{t+1}] \tag{7}$$

$$q_t = U_{2t}(C_{1t},C_{2t})/\ U_{1t}(C_{1t},C_{2t}) \tag{8}$$

Equation (7) simply represents the marginal equivalence of costs and benefits to the consumer for acquiring one additional dollar's worth of credit goods. Equation (8) states the fact that the gross nominal interest rate, $(1+r_t)$ being the inverse of the price of nominal bonds, q_t, is simply the relative price of cash to credit goods.

In equilibrium $M_t = \bar{M}_t$ for all t and assuming that the cash constraint is binding, cash balances are entirely spent on cash goods while the goods market clears. All these mean that :

$$C_{1t} = m_t \tag{9}$$

$$C_{2t} = Y_t - m_t \tag{10}$$

where $m_t = M_t / P_t$ is the equilibrium real balance.

Proposition :

In a stochastic steady state, the relative holding of cash to credit goods is a constant and is given by the following expression:

$$m/(Y-m) = \bar{N}\,[\,E(1/\omega)]^{\xi} \tag{11}$$

where $\bar{N} = [\alpha\beta/(1-\alpha)]^{\xi}\ \sigma^{(1-\gamma)\xi}$ and $\xi = 1/(1+\rho)$

Proof : Appendix.

It is now straightforward to examine the comparative static effects of an increase in uncertainty in the monetary growth rate ω on the relative holding of cash to credit goods. Notice that $1/\omega$ is a convex function of ω and hence its expectation unambiguously increases in response to a mean preserving spread of the distribution of ω, thus raising the relative holding of cash goods. An increase in uncertainty in money growth simply raises the expected rate of return on cash goods which is nothing but E $(1/\omega)$. The consumer, therefore, substitutes away from credit goods to cash goods.[2] The extent of substitution of course depends on the temporal elasticity of substitution between cash goods and credit goods, which is represented by ξ . To see this dependence observe from (11) the useful relation :

$$d\log\,[m/(Y-m)] / d\log E\,(1/\omega) = \xi \tag{12}$$

---

[2] Notice that unlike the standard result about savings under uncertainty [see, for example Leland (1968)], this particular effect of monetary uncertainty on cash holding does not depend on the third derivative of the utility function and it is quite robust to alternative functional forms for the utility function. For example, Basu and Dua (1992) use a nonhomogeneous utility function and obtain the same result.

which means that in equilibrium ξ turns out to be the elasticity of the relative holding of cash to credit goods with respect to a change in the expected rate of return on cash goods.

2.1 Implications for the velocity and the Nominal Interest Rate

The income velocity of circulation is simply Y/m which is inversely related to the relative holding of cash goods to credit goods. The nominal interest rate, $1+r_t$, is given by the reciprocal of the bond price q_t which means that in equilibrium it is represented as :

$$1+r_t = [\alpha / (1-\alpha)] [(Y_t - m_t) / m_t] \tag{13}$$

which is also inversely related to the relative holding of cash to credit goods.

From (11) and (13) it is, therefore, evident that both velocity and the nominal interest rate decline in response to an increase in uncertainty in the monetary growth rate. The magnitude of the velocity and the nominal interest rate [as evident from (12)] is crucially governed by the elasticity of substitution parameter, ξ . The higher the value of ξ , the greater the decline in velocity and the nominal interest rate in response to an increase in monetary growth uncertainty. One polar case is worth mentioning, when $\xi = 0$, in which velocity is invariant to change in the variability of monetary growth rate because the indifference curve between cash and credit goods is simply right angled making substitution between cash and credit goods impossible.

3. Conclusion

This parametric example helps us to understand the relationship between uncertainty in monetary growth rate and velocity. The result shows that velocity and the nominal interest rate unambiguously decline in response to a greater uncertainty in monetary growth rate due to change in the return on holding of cash goods relative to credit goods. The magnitude of the changes in velocity and the nominal interest rate depend in an important way on the elasticity of substitution between cash and credit goods. This exercise also emphasizes that while testing empirically how signigicant the responsiveness of velocity is to changes in variability of money growth, it is important to understand the testable restriction on the deep parameter governing the substitution between cash and credit goods.

Appendix

Derivation of Equation 11 : In equilibrium the first order condition (7) can be rewritten as :

$$m_t . [U_2(m_t, Y_t - m_t) / U_1(m_t, Y_t - m_t)] . U_1(m_t, Y_t - m_t)$$
$$= \beta E_t [(m_{t+1} / \omega_{t+1}) . U_1(m_{t+1}, Y_{t+1} - m_{t+1})] \qquad (A.1)$$

Using the functional form (4) and exploiting the homogeneity property of the utility function, (A. 1) can be written as :

$$m_t^{1-\gamma} [m_t / (Y_t - m_t)]^{1+\rho} \; U_1[1, (Y_t - m_t) / m_t] =$$
$$[\alpha\beta / (1-\alpha)] . E_t[(m_{t+1}^{1-\gamma} / \omega_{t+1}) . U_1\{1, (Y_{t+1} - m_{t+1}) / m_{t+1}\}] \qquad (A.2)$$

Conjecturing a solution: $m_t = \lambda Y_t$ and plugging into (A. 2) one gets :

$$[\lambda / (1-\lambda)]^{1+\rho} = [\alpha\beta / (1-\alpha)] \, \sigma^{1-\gamma} E_t[1 / \omega_{t+1}] \qquad (A.3)$$

Since $\{\omega_t\}$ is an i.i.d. sequence, the right hand side is independent of the conditioning set and hence the method of undetermined coefficient works in this case. From (A. 3) equation (11) immediately follows.

REFERENCES

Basu, P and P, Dua, Velocity Instability : A Monetary or Real Phenomenon, "Working Paper

Friedman, M, "Lessons from the 1979 – 82 Monetary Policy Experiment," American Economic Review, May 1984, 74, 397 – 400

Leland, H, "Savings and Uncertainty : The Precautionary Demand for Saving," Quarterly Journal of Economics, LXXXII (1968), 465 – 73

Lucas, R. E., Jr, and N. L. Stokey, "Money and Interest Rate in a Cash-In-Advance Economy," Econometrica, Vol. 55, No. 3, May 1987, 491 – 513

Welfare Loss from Rent-Seeking for Budgetary Allocation*

Satya R. Chakravarty

1. Introduction

Government restrictions on different economic activities in a market economy generate various forms of rents. Market participants (e.g., firms and individuals) exert effort to achieve such rents and in the process of this effort resources are used up to obtain rents. Thus, wasteful expenses are incurred for securing rents.

Since the publication of Tullock's (1967) pioneering article on rent-seeking behavior, attempts have been made to study the effects of rent-seeking in different contexts. Posner (1975) analysed the case of rent-seeking behavior in a contest between firms to obtain a monopoly position and inferred that all monopoly rents should be counted in the costs of monopoly. [Fisher (1985) provides further discussions on Posner's work.] Krueger (1974) developed a model of competitive rent-seeking for the important case when rents originate from quantitative restrictions on international trade. Katz and Nitzan (1987) studied the problem of rent-seeking in the case of pure public goods. Abbot and Brady (1990) developed a simple model to explore the welfare implications of innovation rent-seeking, obtained by loosening of restrictions on air pollution permit trading. [See Katz and Rosenberg (1989) for further discussions along this line].

Tullock (1967, 1975, 1980) provided several illustrations to demonstrate how lobbying in the corridors of power by different economic agents can give rise to wasteful expenses. He pointed out that this particularly arises in the context of allocation of the government budget to different categories. Katz and Rosenberg (1989) suggested a device for measuring the waste generated due to rent-seeking within the government butget. However, they did not relate their measure to any social welfare function. More precisely, they did not develop their measure as the size of welfare loss resulting from rent-seeking within the government budget. This paper aims at explicity determining the loss of welfare that arises due to rent-seeking behavior in the context of the government budgetary allocation (including transfers).

We begin section 2 with a discussion of the Katz-Rosenberg model. The remainder of the section is devoted to the development of our welfare theoretic approach. Section 3 makes some concluding remarks.

2. Rent-Seeking for Budgetary Allocation and Welfare Loss

Katz and Rosenberg (1989) assumed that the estimates of rent-seeking induced by the government budget are based on two postulates :

(1) A change in the proportional composition of total government spending takes place as a result of rent-seeking activity by different pressure groups.

(2) The total expenditure by different groups to obtain the rent is equal to the total amount of change (in monetary terms) in budgetary allocation.

The first postulate reflects the view that any change in the proportional distribution of the government budget occurs because of lobbying of pressure groups, the groups which exert effort to increse their shares in the budget. The second postulate shows that the rent-seeking behavior under consideration is of competitive character. Consequently, the net aggregate benefit is. zero.

Let x_{ti} be the size of the government budget spent on category i in year t. Suppose that there are n such categories. The budget vector (x_{t1}, x_{t2}, ..., x_{tn}) in year t is denoted by x_t. We write X_t for the size of the government budget in year t, that is, $X_t = \sum_{i=1}^{n} x_{ti}$. The set of all budget vectors is D_n, the non-negative orthant of the Euclidean n-space R^n with the origin deleted. By deleting origin from the set of budget vectors we ensure that the size of the budget is positive. The proportion of the budget spent on category i in year t is $s_{ti} = x_{ti} / X_t$. We write s_t for the vector (s_{t1}, s_{t2}, ..., s_{tn}). According to assumption (1) if the sequence $\{s_{ti}\}$ in year t differs from the corresponding sequence $\{s_{(t-1)i}\}$ in year (t–1), then the difference occurs because of rent-seeking behavior of the pressure groups. More precisely, if the absolute gaps $| s_{ti} - s_{(t-1)i} |$

deviate from zero, then the deviations occur as a result of lobbying of pressure groups.

The first Katz-Rosenberg measure that 'captures rent-seeking for budgetary allocation as a proportion of overall government spending' is

$$Z_t = \frac{1}{2} \sum_{i=1}^{n} \left| s_{ti} - s_{(t-1)\,i} \right|$$

(The division by 2 is done to avoid double counting). Z_t is the sum of absolute changes in the proportions allocated to different categories in year t over year (t–1). Clearly, Z_t is compatible with postulates (1) and (2). Katz and Rosenberg considered the mean value Z_c of Z_t's over time, that is, $Z_c = \sum_{t=1}^{T} Z_t / T$, where T is the number of years considered, as a measure of mean rent-seeking in country c for the government budgetary (and transfer) allocation.

Now, to develop a welfare theoretic measure of rent-seeking in the context of the government budgetary allocation, we consider a social welfare function W, a real valued function defined on D^n. It will rank alternative budgetary allocations in terms of the preference of the society. W is assumed to be continuous, increasing, quasiconcave and homothetic. Continuity of W will ensure that small changes in one or more arguments of W will have little impact on social welfare. Increasingness means that if the size of the budget spent on one category increases, others remaining fixed, then social welfare must increase. Quasiconcavity guarantees that W possesses an indifference curve which is convex to the origin. Homotheticity means that the entire set of indifference curves can be generated by radial expansion or contraction of a single curve (using rays proceeding from the origin). That is, W can be expressed as an ordinal transform of a linearly homogeneous function[1]. Homotheticity is a sufficient condition to guarantee that the size of welfare loss depends on budget shares only. Hence it will not depend on the unit in which x_{ti}'s are measured. That is, the loss function will be free from money illusion - if x_{ti}'s are measured in pounds instead of in dollars its value should not alter.

To ensure that the size of welfare loss should take care of the absolute changes in the budget proportions allocated to different categories between year t and year (t–1), we scale up x_{ti} by $f\ (d_{ti})$, where $d_{ti} = |s_{ti} - s_{(t-1)i}|$ and the function f is continuous, positive, decreasing and $f\ (0) = 1$. Positivity of f is necessary to keep the sign of $x_{ti}\ f(d_{ti})$ same as that of x_{ti}. Decreasingness reflects the view that rent-seeking is socially undesirable. The normalization $f(0) = 1$ guarantees that there will be no welfare loss if $s_{ti} = s_{(t-1)i}$ for all $i = 1, 2, \ldots, n$.

We now define the unchanged share equivalent proportion of budget ρ_t by

$$W\ (\rho_t\, x_t) = W\ [x_t\ f(d_t)] \qquad (2)$$

where $x_t\ f(d_t) = [x_{t1}\ f(d_{t1}), x_{t2}\ f(d_{t2}), \ldots\ldots, x_{tn}\ f(d_{tn})]$. Thus, p_t is a real number which makes the scaled up vector $r_t\, x_t$ socially indifferent to the vector $x_t\ f(d_t)$. Increasingness of W implies that $0 < r_t \leq 1$.The case $p_t = 1$ is attained when there is no change in proportions of budget allocations between years t and (t–1). Given homotheticity of W, $W\ (r_t\, x_t) = g\ [\overline{W}(\rho_t\, x_t)]$, where $\overline{W}$ is linearly homogeneous and g is increasing in its argument. Using homogeneity of $\overline{W}$, we have

$$\rho_t = \frac{\overline{W}\ [x_t\ f(d_t)]}{\overline{W}(x_t)} \qquad (3)$$

As a measure of the size of welfare loss that arises because of existence of rent-seeking for budgetary allocation in year t, we now consider

$$L_t = 1 - \rho_t$$

$$= 1 - \frac{\overline{W}\ [x_t\ f(d_t)]}{\overline{W}(x_t)} \qquad (4)$$

From (2) it becomes quite clear that the scalar ρ_t is the proportionate reduction in the budget sizes of different categories such that the resultant budget vector produces the same level of social welfare as that produced by the actual budget vector, given the changes in the budget share allocations in year t over year (t–1). Therefore $L_t = (1 - \rho_t)$ can be regarded as the proportion of the total budget in year t which, for a given budget allocation, we would be prepared to sacrifice to eliminate all changes in the budget proportions. Decreasingness of f shows that the greater the changes in the proportions allocated to different categories in year t over year (t–1), the higher will the value of L_t. L_t takes on the minimum value zero when there is no rent-seeking, that is, when $d_{ti} = 0$ for all i=1, 2, ..., n. Hence L_t is compatible with postulate (1) of Katz and Rosenberg. Continuity of W ensures that L_t is a continuous function[2].

It is important to note that Katz and Rosenberg (1989) made assumptions (1) and (2) to get a dollar measure of the extent of rent-seeking. The information requirements for empirical implementation of their measure are the budget shares of different categories in different years. Now, using linear homogeneity of $\overline{W}$, we can rewrite L_t in (4) as $1 - \overline{W}(s_t)\, f(d_t) / \overline{W}(s_t)$. Therefore, the same information are sufficient for empirical calculation of our measure L_t, the fraction of the total budget X_t that we could sacrifice to eliminate rent-seeking.

It is now clear that given (4), to every homothetic social welfare function there corresponds a different measure of the size of welfare loss. These measures will differ only in the manner in which we employ our value judgement (through the use of a social welfare function) to aggregate the budget sizes spent on different categories into an overall indicator.

To illustrate the general formula in (4), let us suppose that the social welfare function is the mean of order $\Theta\ (\leqq 1)$. That is,

$$\overline{W}(x_t) = \Big(\sum_{i=1}^{n} a_i x_{ti}^{\Theta}\Big)^{1/\Theta}, \quad \Theta \leqq 1, \quad \Theta \neq 0, \tag{5}$$

$$\prod_{i=1}^{n} x_{ti}^{a_i}, \quad \Theta = 0,$$

where $0 < a_i < 1$ for all i, $\sum_{i=1}^{n} a_i = 1$ and $x_t = (x_{t1}, x_{t2}, \ldots, x_{tn}) \in D^n$.

The requirement $\Theta \leqq 1$ arises because of quasi-concavity. The parameter Θ determines the curvature of the social indifference surfaces. A decrease in the value of Θ will make the social indifference curve more convex to the origin. For $\Theta = 1$, $\overline{W}(x_t)$ becomes linear and as $\Theta \to -\infty$, $\overline{W}(x_t) \to \min_i \{x_{ti}\}$. Assuming that $f(d_{ti}) = 1/(1 + d_{ti})$, we can now determine the explicit form of L_t for the social welfare function in (5). (Note that this functional form of f is taken for an illustration. One can choose a different form obeying the restrictions imposed on f for this purpose. For example, the choice of $f(d_{ti}) = e^{-\alpha d_{ti}}$, where $\alpha > 0$, will also do.)

An alternative of interest arises from the following specification of $\overline{W}$:

$$\overline{W}(x_t) = \frac{1}{b} \sum_{i=1}^{n} b_i \, x_{ti}^{o} \tag{6}$$

where $0 < b_1 \leqq b_2 \leqq \ldots \leqq b_n$, $b = \sum_{i=1}^{n} b_i$ and $x_t^{o} = (x_{t1}^{o}, x_{t2}^{o}, \ldots, x_{tn}^{o})$ is that permutation of x_t such that $x_{t1}^{o} \geqq x_{t2}^{o} \geqq \ldots \geqq x_{tn}^{o}$. Given that x_{ti}^{o}'s are non-increasingly ordered, non-decreasingness of the sequence $\{b_i\}$ is necessary and sufficient for quasi-concavity of the social welfare function in (6) (see Chakravarty, 1990). If we assume that $b_i = (2i - 1)$, then $\overline{W}$ in (6) becomes the Gini social welfare function. Clearly, given (6), we can derive the loss measure L_t in (4) under an appropriate specification of f. Thus, as argued above, alternative forms of L_t will differ only in the manner in which we decide to employ our value judgement.

3. Conclusions

Katz and Rosenberg (1989) suggested a measure of rent-seeking for the government budgetary allocation. The measure is based on the total change in the budget's proportional allocation for different purpose between consecutive years. This paper looks at the size of welfare loss associated with rent-seeking activity for the government budgetary allocation. To every homothetic social welfare function there will be a different measure of the size of welfare loss. These measures explicitly take into account the changes in the proportions allocated to different categories in a particular year over the preceding year.

Notes

* I am grateful to Gordon Tullock for his suggestions.

1. A real valued function H defined on D^n is linearly homogeneous if for all scalars $c > o$ and for all $y \in D^n$, $H(cy) = c\,H(y)$.

2. We can as well measure the loss in welfare from rent-seeking in difference form $W(x_t) - W(x_t\ f(d_t))$, rather than in ratio form, as has been done in (4). But in this case the welfare loss measure will depend on budget sizes of different categories. Hence it will not be free from money illusion.

References

Abbott, A F and Brady, G.L (1990) . 'Innovation-induced rent-seeking the case of air quality management', *Journal of Institutional and Theoretical Economics* 146 : 328-344.

Buchanan, J.M , Tollison, R.D and Tullock, G. (1980) : *Toward a theory of rent-seeking society* Texas A and M University Press, College Station.

Chakravarty, S.R. (1990) . *Ethical Social Index Numbers*, Springer-Verlag, London.

Fisher, F (1985) . 'The social costs of monopoly and regulation : Posner reconsidered' *Journal of Political Economy* 93 : 410-416.

Katz, E. and Nitzan, S (1987) : *Rent-seeking for pure public goods.* York University.

Katz, E. and Rosenberg, J. (1989) : 'Rent-seeking for budgetary allocation : Preliminary results for 20 countries.' *Public Choice* 60 133-144.

Krueger, A.O (1974) : 'The Political economy of the rent-seeking society', *American Economic Review* 64 . 291-303.

Posner, R.A. (1975) 'The social cost of monopoly and regulation', *Journal of Political Economy* 83 : 807-827.

Tullock, G. (1967) : 'The welfare cost of tariffs, monopolies and theft', *Western Economic Journal* 5 : 224-232. Reprinted in Buchanan, Tollison, and Tullock (1980)

Tullock, G. (1975) : 'Competing for aid', *Public Choice* 21 : 41-51.

Tullock, G. (1980) : 'Efficient rent-seeking', In Buchanan, Tollison and Tullock (1980) : 97-112.

Weak* Axiom of Independence and the Non–Expected Utility Theory

Tapan Biswas

1. Introduction

The axiomatic foundation of the expected utility theory (which states that given a set of uncertain prospects individuals pick up the prospect which yields the highest exepected utility) was first laid down by Von Neumann and Morgenstern (1947). Their work inspired various authors to suggest alternative sets of axioms compatible with the expected utility theory. The axiom of independence is one of the axioms which is used, either directly or indirectly, to establish the expected utility theory. This axiom has come under severe criticisms in recent years. A large number of experiments has shown that in making decisions involving uncertain prospects people frequently violate the independence axiom. In this paper, we shall consider the problem of choice under uncertainty from a wider point of view and we shall examine the nature of the restriction imposed by the axiom of independence. We shall use the mean-variance utility function to prove our point. Then we shall consider a weak version of the independence axiom namely the weak* axiom of independence. This is the point of departure from the expected utility theory to the realm of the non-expected utility theory. The weak* axiom allows aversion to pure uncertainty and, in the context of the mean-variance utility theory, it is compatible with utility being an inceasing function of expected returns at all levels.

2. Choice In The Moment-Space Of Lotteries

Consider any finite interval [a, b] on the real line as the outcome (prizes) space for lotteries which may be regarded as probability distribution functions, F (.), defined on the outcome space. We consider the class of lotteries for which all moments of the corresponding distribution functions exist (some of them may be zero) and the distribution functions are uniquely defined by their moments. This uniqueness property requires some minor restrictions on the distribution functions and excludes some pathological cases [see Karlin and Shapley (1953) or Widder

(1949, pp. 29-31, 60)]. We define the i-th non-central moment of a distribution as $\mu_i = E(x^i)$. The utility from an uncertain prospect F (.) may be written as a functional,

$$U = U[F(.)] = U(\mu_1, \mu_2, \mu_3 \ldots\ldots) \qquad (1)$$

Eq. (1) merely states that we can define the utility functions on the moment space of lotteries. We shall see that any compound lottery, G, involving two lotteries F^1 and F^2 with probabilities p and (1–p) respectively may be regarded as a convex combination of F^1 and F^2 in the moment space of the lotteries, i.e., if

$$F^1: [\mu_1(1), \mu_2(1), \ldots\ldots) \text{ and } F^2: U[\mu_1(2), \mu_2(2), \ldots\ldots]$$

then,

$$G: [p.\mu_1(1) + (1-p).\mu_1(2), p.\mu_2(1) + (1-p).\mu_2(2), \ldots\ldots]$$

In order to see this we need the following theorem.

Theorem 1.

Let F^1 and F^2 be two arbitrary lotteries with $\mu_t(1)$ and $\mu_t(2)$ as their non-central moments of order t, then the moment of order t for the compound lottery G = (F^1, p; F^2, 1–p) will be given by

$$\mu_t(G) = p.\mu_t(1) + (1-p).\mu_t(2)$$

Proof

Let f_1, f_2 and g be the probability density functions of $F^1(x_1)$, $F^2(x_2)$ and G(x) respectively. Remember, G is a compound lottery yielding either of the lotteries F^1 and F^2 as outcome. G can yield a final outcome, say d, either through F^1 achieving d or through F^2 achieving d. Therefore,

$$\text{Prob}\,(x = d \pm \epsilon) = p.[\text{Prob}\,(x_1 = d \pm \epsilon)] + (1-p).[\text{Prob}\,(x_2 = d \pm \epsilon)]$$

This implies that the probability density function for the compound lottery G is given by : $F_G(x) = p.\ F_1(x) + (1–p).\ F_2(x)$. Hence the non central moment of G for any order t is given by,

$$\mu_t(G) = \int x^t.f_G(x) = \int x^t.[\,p.\ f_1(x) + (1-p).f_2(x)\,].dx$$
$$= p.\ \mu_t(1) + (1-p).\mu_t(2)$$

(Q. E. D.)

Henceforth, for the sake of avoiding unnecessary repetitions the non-central moments of order t will be referred to simply as moments and denoted by μ_t.

Let us now turn to the independence axiom and a weaker version of it called the betweenness axiom. We shall explain how they restrict the form of the utility function in Eq. (1). Consider, three lotteries F^1, F^2 and F^3. Construct two compound lotteries G^1 and G^2 such that,

$$\begin{aligned} G^1 &= (F^1, p \,;\, F^3, 1-p) \\ G^2 &= (F^2, p \,;\, F^3, 1-p) \end{aligned} \tag{2}$$

G^1 is a compound lottery with p as the probability of winning F^1 and 1–p as the probability of winning F^2. G^2 should be interpreted in a similar way. Following the convention used by Sen (1970) we use the symbols P, R and I to denote strict preference, weak preference and equivalence respectively. For example, F^1 R F^2 means F^1 is weakly preferred to F^2.

Axiom of Independence : F^1 P F^2 (or, F^1 I F^2) implies G^1 P G^2 (or, G^1 I G^2).

Betweenness Axiom : F^1 I F^3, implies F^1 I F^3 I G^1.

The independence axiom serves as a corner stone of the expected utility theory. In an effort to generalize the expected utility theory, some authors proposed the betweenness axiom which is weaker than the independence axiom. Let F^3 be a compound lottery (F^1, p ; F^2, 1–p) where F^1 I F^2 [i.e. U (F^1) = U (F^2)]. Theorem 1 together with the betweenness axiom imply that, in the moment space of lotteries, F^3 is a convex combination of F^1 and F^2 for any value of p. In other words, betweenness axiom implies that the indifference curves in the moment space of lotteries are linear. Since the betweenness axiom is weaker than the independence axiom, it follows that the linearity property of the indifference curves also holds under the expected utility theory. In fact, the expected utility theory has a stronger implication, namely, the functional U [F (.)] in Eq. (1) is linear.

Let us consider an example of what we have discussed so far. Suppose, for an individual seeking to maximize his utility, only the first two moments of the distribution matters, i.e., for him $U = U(\mu_1, \mu_2)$. The reader should note that this is the mean-variance utility theory because of the relationship, $\sigma^2 = \mu_2 - \mu_1^2$. According to the expected utility theory (independence axiom), U is linear. Let $U = a + b\mu_1 - c\mu_2$. Then we have

$$U = a + b.\mu_1 - c.\mu_1^2 - c.\sigma^2 \qquad (3)$$

The mean-variance utility theory normally asumes that individuals are risk averse. For this reason we have a negative coefficient for σ^2. From Eq. (3), it is clear that U is not increasing in μ_1. After a certain value of μ_1, any increase in μ_1 will lead to a decline in U. This is quite unacceptable for lotteries involving financial gains. Any increase in the expected returns from the lotteries is always a good news. The reason for this decline in utility is obvious. Under the expected utility theory, mean-variance utility implies that (i) either the lotteries are restricted to only normal distributions or (ii) the utility function is quardratic. We are considering a general class of distributions because a compound lottery formed out of normal distributions is, in general, not a normal distribution. A quadratic utility function, $U = a + bx - cx^2$ implies $U = a + b\mu_1 - c\mu_2$ as above. It also implies that after a certain level of income, utility is a declining function of income, x.

Under the betweenness axiom, we have

$$U = a(U) + b(U).\mu_1 - c(U).\mu_2 \qquad (4)$$

In this case, the indifference curves are linear in μ_1 and μ_2. However, in order to avoid any intersection of the indifference curves, another axiom called the Weak Axiom of Independence [Chew and MacCrimmon (1979)] is often invoked. It can be shown that even under betweenness, utility is not monotonically increasing in μ_1 [Biswas (1994)]. The argument is simple. Since, $\sigma^2 = \mu_2 - \mu_1^2$, the combinations of μ_1 and μ_2 which satisfy the zero variance condition ($\sigma^2 = 0$) is given by $\mu_1 = \mu_2^{\frac{1}{2}}$. When the indifference curves in the (μ_1, μ_2) space are linear, they must twice intersect the curve satisfying the relationship $\mu_1 = \mu_2^{\frac{1}{2}}$. Therefore, there are two values of μ_1 with $\sigma^2 = 0$ which gives the same level of utility. This means $U(\mu_1 \mid \sigma^2 = 0)$ reaches a maximum for some value of μ_1.

3. The Weak* Axiom Of Independence

We have seen that linear indifference curves in the moment space of lotteries have an undesirable implication. Allais (1984) argued that the expected utility theory (more precisely the independence axiom) suffers from a serious drawback. Suppose, utility u(x) ia an increasing function of x. A lottery defined on $x \in [a, b]$ also induces lottery defined on the utility space, $u \in U$, because of the monotonic relationship between u and x. In the expected utility theory, the axioms lead to the conclusion that in comparing lotteries, only the expected utility is taken into account, whereas according to Allais, other moments of the utility distribution may play a crucial role in the choice. It is this idea which, led him to construct the famous "Allais paradox". For Allais' own formulation of a non-expected utility theory, the reader is referred to Allais (1984).

Earlier in the paper we have referred to the axiom of weak independence which runs as follows.

Axiom of Weak Independence : Consider, three different lotteries F^1, F^2 and F^3, such that $F^1 \; I \; F^2$. Given any probability 'p', there exists a probability 'q' for all possible F^3, such that $L_1 \; I \; L_2$ where $L_1 = (F^1, p \, ; \, F^3, 1-p)$ and $L_2 = (F^2, p \, ; \, F^3, 1-p)$.

The weak axiom of independence is weaker than the betweenness axiom but it also requires the indifference curves in the moments space of lotteries to be linear and precludes any intersection of linear indifference curves in the region of feasibility. This is not good enough if we want utility to increase with increase in certain income, x, i.e., $U(\mu_1 \mid \sigma^2 = 0)$ increasing positively with μ_1. With this in mind we state a weaker version of the independence axiom.

Weak* Axiom of Independence : Consider three lotteries F^1, F^2 and F^3 such that F^1 I F^2, F^1 P F^3 and F^2 P F^3. We assume that given any probability 'p', there exists a probability 'q' such that L_1 I L_2 where $L_1 = (F^1, p ; F^3, 1-p)$ is pereferred to F^3 and $L_2 = (F^2, q ; F^3, 1-q)$.

The weak* axiom of independence, henceforth called the weak* axiom, rules out the possibility of strictly convex preferences. Let μ denote $(\mu_1, \mu_2, \ldots)$ and consider two arbitray vectors $\mu(1)$ and $\mu(2)$ such that $U[\mu(1)] = U[\mu(2)]$. Preferences are said to be non-convex if

$$U\left[t.\mu(1) + (1-t).\mu(2)\right] \leq U\left[\mu(1)\right] ; 0<t<1$$

In case of mean-variance utility, non-convex preferences imply that the indifference curves are weakly concave in the $\mu_1 - \mu_2$ plane (see figure 1). Note, weak concavity includes the possibility of linear indifference curves.

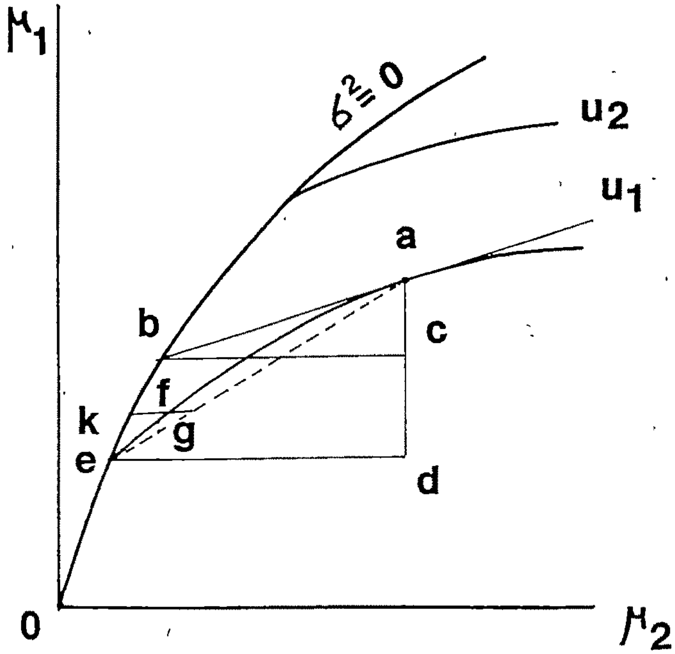


Figure 1

Theorem 2

If the weak* axiom of independence holds and $U(\mu)$ is continous in μ, then preferences must be non-convex everywhere in the moments space.

Proof

If preferences are not non-convex everywhere, and U is continuous, there must exist a convex subspace D of the moments space within which the preferences are convex. Consider, $\mu(1)$, $\mu(2)$, $\mu(3) \in D$ such that $U[\mu(1)] = U[\mu(2)]$ and $\mu(3) = \mu(1) + (1+\delta)[\mu(2) - \mu(1)]$. If μ has a finite dimension, $\mu(3)$ lies on the line joining $\mu(1)$ and $\mu(2)$ and is slightly outside $\mu(2)$. Convexity of preferences imply that $U[\mu(2)] > U[\mu(3)]$. Consider, $\mu(4) = p.\mu(1) + (1-p).\ \mu(3)$ where p is close to 1 so that due to the convexity of preferences

$$U[\mu(4)] > U[\mu(1)] = U[\mu(2)].$$

By the weak* axiom of independence, there must exist a 'q' such that $\mu(5) = q.\mu(2) + (1-q).\ \mu(3)$ and $U[\mu(4)] = U[\mu(5)]$. If the preferences are convex,

$$U[\mu(2)] > U[\mu(4)] = U[\mu(5)]$$

because $\mu(2)$ can be expressed as a convex combination of two indifferent vectors of moments, $\mu(4)$ and $\mu(5)$. Thus we arrive at a contradiction with respect to the relationship between $\mu(2)$ and $\mu(4)$, if we allow for the convexity of preferences.

(Q. E. D.)

We must note that the weak* axiom of independence is closely related to an aversion of uncertainty. Even if an individual is indifferent between two lotteries, a compound lottery constructed out of the two, adds to the uncertainty by adding another step to the resolution process. According to the weak*axiom,any unnecessary complexity to the process of resolution may not be desirable and this is reflected in the non-convexity (including possible non-linearity) of preferences.

4. Weak* Axiom, Non-Expected Utility And The Mean-Variance Theory

We shall now discuss the implications of the mean variance utility theory and the role of the weak* axiom in that context. The mean variance utility theory which forms the corner-stone of portfolio analysis assumes that utility is an increasing function of expected returns and is a declining function of the variance,

$$U = U\left(\mu_1, \sigma^2\right) ; \; U_1 > 0, \; U_2 < 0 \qquad (5)$$

It can be shown that if $U_1 > 0$ for all values of μ_1, and U_2 is negative for all values σ^2, then the indifference curves in the $\mu_1 - \mu_2$ plane are concave, i.e., μ_1 is a concave function of μ_2 [Biswas (1994)]. For example, in portfolio theory, it is often assumed that U is a linear function in μ_1 and σ^2 so that the indifference relationships are given by, $\mu_1 = \alpha + \beta . \sigma^2$. This can be simplified as

$$\mu_1 + \beta . \mu_1^2 = \alpha + \beta . \mu_2 .$$

It can be easily checked that $d_{\mu_1} / d_{\mu_2} > 0$ and $d^2\mu_1 / d\mu_2^2 < 0$. We have seen that the independence axiom does not allow for such a relationship and requires that along any indifference curve, μ_1 must be a linear function of the form $\mu_1 = \alpha + \beta . \mu_2$. On the other hand, weak* axiom allows for concave indifference curves by theorem 2. Therefore Eq. (5) is plausible under the weak* axiom of independence.

Figure 1 illustrates the shape of indiference curves under the mean variance utility, satisfying the weak* axiom of independence. Horizontal distance from the curve representing, $\sigma^2 = 0 \left(\text{i.e. } \mu_1 = \mu_2^{1/2}\right)$ equals the variance of the lottery given the value of μ_1. For example, for the lottery represented by the point d, in figure 1, the variance is the line segment, ed. This may be shown easily. Since $\sigma^2 = \mu_1 - \mu_1^{1/2}$, for a fixed value of

μ_1, $\Delta\sigma^2 = \Delta\mu_2$. Between the points e and d, μ_1 is same. Therefore, $\sigma^2 = ed$. The marginal rate of substitution between expected returns and the risk (variance) at 'a' is the slope of the indifference curve at 'a'.

What is the economic significance of concave indifference curves in the $\mu_1 - \mu_2$ plane? Suppose a risk averse individual is considering a purchase of stocks with different variances of returns. Since he is risk averse, he will require higher expected returns from stocks with higher variances to keep him at the same level of utility. But the rate of compensation he will seek for additional risk will decline with the level of risk he has already undertaken compensated by higher expected returns. There is no irrationality in such behaviour, although this is repudiated by the independence axiom. It is also easy to see why such behaviour is consistent with an aversion to pure uncertainty, i.e., a compound lottery composed of two indifferent alternatives is inferior to the alternatives. In figure 1 consider a certain prospect, e, and an indifferent uncertain prospect, a. A compound lottery, g, reduces the expected return and the variance proportionately. However, given the expected return at g, the variance kg is larger than the compensating variance kf. Clearly a proportionate fall in the variance is not enough to keep the individual on the same indifference curve. From figure 1, it is also clear that the individual's risk premium at 'a' is given by the vertical line segment, ad. Note, the concavity of the indifference curves is related to the extent of aversion to pure uncertainty, i.e., aversion to compound lotteries composed of equivalent prospects. Therefore, we would expect the risk premium under the mean variance utility to be at least as large as under the expected utility theory. In figure 1, ac measures the risk premium under expected utility theory. In this case, the indifference curves are linear and premiums decline proportionately with respect to the risk. We call ac the *pure risk premium.* We shall call the difference between the risk premium and the pure risk premium, cd, the *uncertainty premium.* If the individual is indifferent between a compound lottery composed of equivalent prospects and the prospects themselves, then the uncertainty premium vanishes and the risk premium equals the pure risk premium. Also note, when the variance is very small (risk in the small), uncertainty premium is of second order importance and the pure risk premium approaximately equals the risk premium.

5. Conclusion

As we have seen, the weak* axiom of independence, in the context of mean variance analysis, allows aversion to pure uncertainty. But aversion to pure uncertainty need not be true in all cases. "The Doctor's Dilemma" [Sen (1986)] is an example for such a case. Suppose a doctor in the war front has two emergency patients requiring the same operation. But he has enough medical provision for the operation of only one patient.The doctor is indifferent in his choice between the two patients. The doctor can choose the patient directly and be certain of the patient to be operated upon. But he may prefer the use of a lottery of some kind for selecting the patient and thus prefer the introduction of uncertainty in the selection process. Whether a risk- averse individual is also averse to pure uncertainty or not will depend upon the context of the uncertain prospect within which he has to take a decision. Consequently, it is not likely that a simple general theory of choice under uncertainty would ever be found out. The independence axiom is an axiom which neutralizes the effect of pure uncertainty in decisions under uncertainty and simplifies the process of decision making. Unfortunately, numerous experiments have shown that the axiom fails in majority of cases and has generated a large literature on non-expected utility theories.

REFERENCES

Allais M., The Foundations of the Theory of Utility and Risk: Some Central Points of the Discussions at the Oslo Conference In *Progress in Utility and Risk Theory* (O.Hagen and F Wenstop, eds.), (Reidel Publishing Company: Holland), 1984.

Biswas T., Mean–Variance Utility and the Axioms of Independence, prepared for presentation at FUR VII conference, 1994.

Chew S H. and MacCrimmon, Alpha-nu Choice Theory: An Axiomatization of Expected Utility, University of British Columbia, Faculty of Commerce Working Paper No 669, 1979.

Karlin S. and L.S. Shapley, Geometry of Moment Spaces, (Memoirs of the American Mathematical Society, Vol.12), American Mathematical Society, 1953.

Pryce J.D., Basic Methods of Linear Functional Analysis, (Hutchinson· London), 1973.

Von Neumann J. and O. Morgenstern, Theory of Games and Economic Behaviour, (Princeton University Press: Princeton), 1947

Sen A , Rationality and Uncertainty in *Recent Developments in the Foundations of Utility and Risk Theory* (L Daboni, A. Montesano and M Lines, eds.), (Reidel Publishing Company), 1986

Sen A , Collective Choice and Social Welfare, (Holden-Day Inc· London), 1970.

Share Tenancy and Family Resources

Jaydip Datta and Rabindranath Mukhopadhyay

Introduction

How the leasing activity gets motivated by the imperfections in the input-markets and sometime by non-marketability of certain inputs, has been a topic of recent research interest. Bardhan (1976) is an attempt to demonstrate the connection between imperfections in capital market and market for irrigational facilities and market for chemical fertiliser and leased-in cultivation. Bell (1976) is a pioneering study [for the Purnea district of Bihar] to findout how the motivation for lease-in tenancy gets influenced by the fact that there is no market for bullock services. There it has been demonstrated that lease-in decision of a landless household has a direct connection with its endowment of bullocks. Similar was the study of Bliss and Stern (1982) [for Palanpur district of Uttar Pradesh]. Pant (1983) has demonstrated how lease-in decision of a landless tenant gets motivated by the fact that there does not exist any market for farm management services and labour market is imperfect. It is an exercise to show how the lease-in decision of a landless household gets changed due to change in the size and composition of the family when a section of the household labour is unable to earn wage income and entrepreneurial service is an exclusively in-house facility.

The present study provides a framework having close resemblence with Pant (1983) model; it proposes to reexamine the results of the earlier paper. The treatment of managerial service in the present excercise is different from that of Pant (1983); segment of family labour disqualified otherwise to earn wage income is also treated here in a slightly different fashion.

Farm management is a service for which there does not exist any market; i.e., if any household does possess it by virtue of experience, it has it, otherwise not. Another feature, very much prevalent in rural India is that there is a caste barrier–upper caste people do not work on land owned by lower caste people. Again there are caste or social barriers against women members working in others' land. Similarly, women members, child members and old members are physically less capable compared to their able bodied male counterparts as far as strenuous cultivating activities are concerned. Naturally, these members, even if they are allowed to work in others' land, do find themselves discriminated against in wage payment. But these members in their own cultivation can assist or participate in other routine operations involving less physical vigour, namely, weeding, planting or tending to farm animals. Therefore, households efficient in farm management will like to lease-in more land. Similarly, households having large number of female, child or old members will prefer to lease-in more land so that these members of the family can be gainfully employed in own cultivation. As far as these two motivations for lease-in are concerned, the current exercise is in full agreement with Pant (1983).

In Pant (1983) initially managerial labour hours was introduced as a variable in the production function. But very shortly with the help of a conversion factor b >1, those managerial labour hours were converted into ordinary labour hours. This has spoiled, in the opinion of the current paper, the very purpose of having a seperate identity for managerial efficiency. The only concern of Pant (1983) about managerial efficiency comes in the form of - less is the value of b, better is the quality of labour. Moreover, it has considered managerial efficiency in the form of managerial labour hours. This means managerial efficiency does not have much scope to poke its nose. Managerial capability can bring-in two types of differences between households–first between 'haves' and 'have nots' and finally within the 'haves'. To differentiate within the 'haves', it is this question of managerial efficiency which comes in. This managerial efficiency aspect, in the opinion of the present paper, has not been properly tackled in Pant (1983). What is more surprising is that at a later stage there is a further simplified model in Pant (1983) where this variable of managerial labour hours has been dropped. This has necessitated a reconsideration, in the proper perspective, of the aspect of in-house availability of managerial labour hours as far as it acts as a motivation for leasing activity. Accordingly, this has further necessitated to redefine the division of cultivating responsibilities within the household labour force.

The Model

The prime objective of the model is to demonstrate how the lease-in decision of a landless household gets influenced by changes in its family endowment, particularly those elements which are either not marketed or marketed under imperfect conditions; focuss of the model is thus on the behaviour of the landless household. Towards consideration of supply side, the case of landowing household will come at a later stage. The picture will come to a conclusion after the treatment of land market and labour market. The notations and assumptions of Pant (1983) have been retained as far as possible to make the results compatible.

The landless household : By definition it has no land of its own. It organises cultivation on leased-in land with the help of family labour. For cultivation it does not hire any labour; rather labour from this family goes to the wage market to earn wage income as and when it is possible. Child labour, female labour and old members of the family are unable to go to the wage market either due to caste or social barriers or due to market imperfections; thus they constitute the dependent labour force of the family. Rest of the members of the family are capable of earning wage income; thus they are identified as the wage employable labour of the family. Therefore, if L stands for the total labour endowment of the landless household, it can be expressed as $L = \bar{l} + \bar{\bar{l}}$ where $\bar{l}$ represents dependent labour and $\bar{\bar{l}}$ denotes wage employable labour of the said household, who can also work for own cultivation.

The entire gamut of cultivating activities is broadly divided into three parts–activities requiring strenuous physical labour, like land preparation, routine activities and managerial activities. Dependent labours can perform only routine works; rest of the works are carried out by wage employable labour. But it needs to be mentioned that wage employable labour can perform routine works also. It is asumed that labour requirement per unit of land preparation is fixed.[1] If 'a' stands for the labour requirement per unit of land preparation, then 'ah' is the labour required to prepare n units of leased-in land.[2] If m represents total supervisory labour hours devoted on lesed-in cultivation, then l_1 can denote part of the wage employable labour engaged for non-routine activities on leased-in cultivation, when $l_1 = ah + m$. Accordingly, let l_2 represent household labour engaged in routine operation on leased-in cultivation; $l_2 > l$ will imply $l_2 - \bar{l}$ units of wage employable labour has been engaged into routine operation. Finally, $L - l_1 - \left(l_2 - \bar{l}\right)$ will represent units of household labour that has gone to earn wage income.

It is assumed that there does not exist market for farm management services. This means that the household responsible for organising cultivation will have to provide this managerial service on its own either on leased-in land or on own land. Inter-household differences in the in-house availability of managerial capabilities are likely to exist–experienced families are likely to be more efficient in managerial functions. The variable m has been reported to denote total supervisory labour hours of the landless household; thus m can be designed as $m \frac{1}{\alpha} \frac{\bar{l}}{h}$. This means that more is the availability of dependent labour per unit of leased-in land, less is the requirement of managerial supervision.[3]

In other words, this means that $m \propto \frac{h}{l}$ or $m = K.\frac{h}{l}$ where K is the constant of proportionality. This K can be interpreted as the managerial efficiency parameter differentiating efficient households from their inefficient counterparts. If the ratio $\frac{h}{l}$ remains fixed, then every fall in m will lead to a fall in K. Since a fall in m for constant $\frac{h}{l}$ ratio implies increase in supervisory efficiency, K may be interpreted as the managerial efficiency parameter. Thus every fall in K indicates increase in managerial efficiency. Therefore, for any household, less is the value of K, the more efficient is the household[4] This is because, any drop in the value of K implies that the said household is capable to cultivate either more leased-in land with same dependent labour or same leased-in land with less dependent labour in association with same managerial effort. As a result, the present study is more concerned with managerial efficiency parameter K than with managerial labour hours m; Pant (1983) was concerned with the variable m.

The production function confronting the landless household is described as $Q = F(h, l_2; K)$. This means that the said household organises cultivation on h units of lease-in land, applies l_2 units of labour hours for routine cultivating activity and with the help of K degree of managerial efficiency is able to produce Q units of output (grain). Diminishing returns to scale is applicable to F with respect to both h and l_2.

To prepare h units of leased-in land, the household has to devote ah units of labours; this will come from the wage employable labour force of the household. In addition, l_2 units of labour hours are engaged for routine cultivating activity on leased-in land; this covers the entire dependent labour force $\bar{l}$ of the household. And finally, it applies m units of managerial labour hours, which again comes from the pool of wage employable labour force of the said household. Therefore, units of labour hours left with the household that can go to earn wage income can be expressed as $\bar{\bar{l}} - (l_2 - \bar{l}) - ah - m = \bar{\bar{l}} - l_2 + \bar{l} - ah - m$. Let the variables r and w stand for the rental rate and wage rate prevailing in the land market and labour market respectively. Both these variables are endogenously determined but it is assumed that neither the landless household nor the landowning household is capable enough to manipulate them. That means r and w though endogenously determined, are assumed to be given to both the households.

Therefore, corresponding to every wage rate w and rental rate r, income equation of the landless household can be expressed as

$$Y_t = (1-r)\, F(h,\, l_2;\, K) + w[\bar{\bar{l}} - l_2 + \bar{l} - ah - \frac{Kh}{\bar{l}}] \qquad (1)$$

where Y_t denotes income of the tenant household. It is natural to believe that every tenant household will be interested in the following optimising problem:

$$\max_{h, l_2} Y_t \text{ s.t. } l_2 - \bar{l} \geq 0 \text{ and } \bar{\bar{l}} \geq l_2 - \bar{l} + ah + \frac{Kh}{\bar{l}} \qquad (2)$$

Following three different situations emerge from the above optimising problem :

1. When the first constraint is binding i.e., $l_2 - \bar{l} = 0$, it means that the household utilises for routine activity only the dependent labour hours and nothing more than that. No wage employable labour is thus wasted. This leaves the family with the opportunity to send maximum labour to the wage market. The household can thus be identified as the predominantly wage earning family;

2. If the second constraint is binding i.e., $\bar{\bar{l}} = l_2 - \bar{l} + ah + \frac{Kh}{\bar{l}}$, this means that all wage employable labour available with the household is engaged for cultivation on leased-in land and the family is left with no labour for wage income. It is thus a pure tenant household; and finally

3. When neither of the constraints is binding, it means that the household derives income from both ends - cultivation on leased-in land and sending labour to wage market. The household is thus of a mixed status - tenant-cum-wage earner.

Let the third case be the starting point for the present analysis, since this one is likely to represent the most common reality. Since neither of the constraints are binding, problem (2) for the time being is an unconstrained optimising problem. Assuming, that the interior solution exists for such a problem, the first order equilibrium conditions[5] obtained are

$$\begin{aligned} &(1-r)\, F_h(h,\, l_2;\, K) - aw - \frac{wK}{\bar{l}} = 0 \text{ and} \\ &(1-r)\, F_{l_2}(h,\, l_2;\, K) - w = 0 \end{aligned} \tag{3}$$

Thus h and l_2 are solvable in terms of endogenous r, w, K and exogenously given values of a and $\bar{l}$. Let the equilibrium values be denoted by h^* and l^*_2. Considering F to be of Cobb-Douglas variety with diminishing returns to scale in h and l_2, following results are obtained :

$$\begin{aligned} &(a)\ \delta h^* / \delta \bar{l} > 0 \quad (b)\ \delta h^* / \delta w < 0 \\ &(c)\ \delta h^* / \delta a < 0 \quad (d)\ \delta h^* / \delta K < 0 \end{aligned} \tag{4}$$

From the result (4a) it follows that for every increase in the dependent labour force of the landless household, there has to be an increase in the leased-in land. This speaks about the effect of change in family composition on lease-in activity. Pant (1983) was of the opinion that change in the family composition of the tenant household will not have any effect on its lease-in decision.

Result (4b) speaks about the effect of change in the wage rate on lease-in tendency : corresponding to every increase in the wage rate there will be a fall in the lease-in demand of the landless household. Surprisingly, Pant (1983) had the opposite result : tenancy will be more in regions with higher wage rate.

Result (4c) implies that in regions with high rainfall, meaning thereby that land preparation is relatively cheap because land is more fertile, there will be an increased demand for leased-in land from the tenant household. Here again a surprising encounter with Pant (1983) : tenancy will be more in regions with less rainfall.

Finally, result (4c) gives the idea that the more efficient the household is in farm management, the higher is the demand for leased-in land from the said family (reference may be made to footnote 4).

It is now time to switch over to the two extreme cases, namely, the case of a pure tenant and the case of a predominantly wage earner. In problem (2) when the first constraint is binding i.e., $l_2-\bar{l}=0$, it means that the landless household is fortunate enough that it does not have to waste any wage employable labour for routine cultivating activity on leased-in land. Therefore, the optimising problem (2) now takes the shape of

$$\max_{h,l_2} Y_t \text{ s.t. } l_2-\bar{l}=0$$

Assuming that interior solution exists for such a problem, the first order equilibrium conditions[6] obtained are

$$\begin{aligned} &(1-r)\, F_h(h,l_2;K)-aw-\frac{Kw}{\bar{l}}=0 \\ &(1-r)\, F_{l_2}(h,l_2;K)-w-\lambda \quad =0 \qquad (5) \\ &\bar{l}-l_2 \qquad\qquad\qquad\quad =0 \end{aligned}$$

where λ is the Lagrange multiplier. From the system (5) of three equations, equilibrium values h^*, l^*_2 and λ^* are obtained in terms of endogenous K, r, w and exogenously provided values of a and $\bar{l}$. Again, assuming F to be obeying diminishing returns to scale in h and l_2, it can be shown that[7]

$$\delta h^* / \delta \bar{l} > 0 \tag{6}$$

This means that every increase in the dependent labour force of the household will lead to an increased demand for leased-in land. This result bears harmony with its counterpart (Pant, 1983). Further, results (6) and (4a) speak in the same voice. Thus effect of change in the composition of the labour force of the landless household on leased-in demand will be same, whether it is a predominantly wage earning household or a household deriving income from both wage market and tenant cultivation.

Turning to the case of a pure tenant [i.e., after organising cultivation on the leased-in land, the landless household is left with no surplus wage employable labour that can bring wage income],the optimising problem (2) now takes the form of

$$\max_{h,l_2} Y_t \text{ s.t. } \bar{\bar{l}} = l_2 - \bar{l} + ah + \frac{Kh}{\bar{l}}$$

Assuming that interior solution exists for such a problem, first order equilibrium conditions[8] obtained are

$$\begin{aligned} &(1-r)\, F_h(h, l_2 : K) - a(w+\gamma) - \tfrac{K}{\bar{l}}(w+\gamma) = 0 \\ &(1-r)\, F_{l_2}(h, l_2 ; K) - (w+\gamma) \qquad = 0 \\ &\bar{\bar{l}} - l_2 + \bar{l} - ah - \frac{Kh}{\bar{l}} \qquad = 0 \end{aligned} \tag{7}$$

where γ is the Lagrange multiplier. Since (7) represents a system of three equations with three variables h, l_2 and γ, it is a determinate system and accordingly equilibrium values h^*, l_2^* and γ^* are obtained in terms of endogenous variables, r, w, K and exogenously provided values of a and $\bar{l}$. Assuming F to obey diminishing returns to scale in h and l_2, the following result[9] can be demonstrated to be true :

$$\frac{\delta h^*}{\delta \bar{l}} > 0 \tag{8}$$

Thus again, it is revealed that with every increase in the dependent labour force of the landless household, there will be increase in the demand for leased-in land. Therefore, in all the three cases – pure tenant, predominantly wage earner and a combination of the two – the landless household is homogenous in its attitude towards lease-in activity : for every increase in the dependent labour force within the family, there will be an increase in lease-in demand from the household; and opposite otherwise. This speaks about the effect of change in the composition of the labour force of the household on its leased-in demand. Pant (1983) had similar results with respect to both pure tenant household and predominantly wage earning household; these were obtained under a simplified model after dropping from explicit consideration the managerial role of the labour. However, when the landless household is simultaneously a wage earner and a tenant cultivator, Pant (1983) has demonstrated that its lease-in decision is independent from the size and composition of the household labour force; this is true in both the general model and the simplified model; possibly this is due to over simplification of the model (Pant, 1983).

The landowing household : It leases out q units of land to the tenant and organises cultivation on $1 - q$ units of land with the help of hired labour. It is assumed that members of the landowning household do not take part in cultivating activity except performing supervisory function. Regarding explicit consideration of the managerial role of the labour force from the landowning household, following the definition of K in the case of landless household, it is defined as $\bar{K} = \bar{m}[l_3/(1-q)]$ where l_3 is the labour hours hired for cultivation[10] and $\bar{m}$ varies inversely with labour-land ratio $[l_3/(1-q)]$. Accordingly, $\bar{K}$ is the managerial efficiency parameter of the landowning household. The income equation of such a landowning household can be expressed as

$$Y_l = rF(q, l_2, K) + G(1-q, l_3, \bar{K}) - wl_3 \tag{9}$$

where Y_l denotes income of the landowning household, G stands for the production function of the landowning household[11] exhibiting diminishing returns to scale in both land and labour, w is the wage rate and r is the rental rate. The optimising problem of the landowning household thus stands out as

$$\max_{q,l_3} Y_l \text{ s.t. } q \geq 0 \text{ and } l_3 \geq 0 \tag{10}$$

Assuming interior solution exists for such a problem, the first order equilibrium conditions[12] obtained are

$$\begin{aligned} r F_q(q,l_2,K)+G_q(1-q,l_3;\overline{K}) &= 0 \\ r F_{l_3}(q,l_2;K)+G_{l_3}(1-q,l_3;\overline{K}) - w &= 0 \end{aligned} \tag{11}$$

Therefore, (11) is a system of two equations in two unknowns q and l_3[13]. Hence equilibrium values q^* and l^*_3 are obtainable in terms of endogenous values of r, w, K, $\overline{K}$. Assuming both F and G to obey diminishing returns to scale in land and labour following results can be demonstrated to be valid :

$$\begin{aligned} &(a)\ \frac{\delta(1-q^*)}{\delta\overline{K}} < 0 \Rightarrow \frac{\delta q^*}{\delta\overline{K}} > 0 \qquad (b)\ \frac{\delta l^*_3}{\delta w} < 0 \\ &\text{and } (c)\ \frac{\delta q^*}{\delta w} > 0 \Rightarrow \frac{\delta(1-q^*)}{\delta w} < 0 \end{aligned} \tag{12}$$

From the result (12a) it follows that with every increase in managerial efficiency (denoted by fall in $\overline{K}$) of the landowing household, it will be interested to organise cultivation on a larger scale and accordingly less will be its tendency to lease out land. Result (12b) implies that as wage rate increases, labour becomes costly, making cultivation organised by hired labour a costly affair and accordingly the landowning household will be less interested in cultivation. The very same implication also follows from the result (12c) – with every increase in wage rate, less will be cultivation organised by landowning household and accordingly more will be the leasing out of land.

Land market and labour market : The description of the model is complete once the determination of equilibrium values of

r and w is made. If it is assumed that leasing operation takes place in the land market under perfect competition, it means that both the landless household and the landowning household are price takers. This in turn implies that lease-in and lease-out operations will take place only at that rental rate which can bring in equality between these two sides. Thus equilibrium value r* is obtained from

$$h^*(r|\bar{l},\bar{\bar{l}}, a) = q^*(r|\bar{l},\bar{\bar{l}}, a) \tag{13}$$

Similarly, if perfect competition is assumed to prevail in the labour market, this will imply that hiring of labour by the landowning household and supply of wage employable labour to the market by the landless household will take place only at that wage rate which bring in equality between demand and supply. Further, neither of the parties is capable to influence the wage rate. Thus equilibrium value w* of the wage rate is determined from

$$\overset{*}{l}_4(w|\bar{l},\bar{\bar{l}}, a) = \overset{*}{l}_3(w|\bar{l},\bar{\bar{l}}, a) \tag{14}$$

$$\text{where } \overset{*}{l}_4 = \bar{\bar{l}} - \overset{*}{l}_2 + \bar{l} - ah^* - \frac{Kh^*}{\bar{l}} \tag{15}$$

Here $\overset{*}{l}_4$ stands for the equilibrium value of wage employable labour sent to the market by the landles household. Once equilibrium values of leased-in land h^* and labour hours engaged in routine cultivating operation $\overset{*}{l}_2$ are determined, relation (15) states that corresponding to given endowment of dependent labour $\bar{l}$ and wage employable labour $\bar{\bar{l}}$, equilibrium supply of labour $\overset{*}{l}_4$ to the wage market is determined.

Conclusion

This paper has thus demonstrated how part of the family resources which are either not marketed or marketed under imperfect conditions, can influence the leasing decision of the household concerned. It has been established that a landless household when organises cultivation as share cropper, will always want to lease-in more land whenever there is increase in the dependent labour force within the family; and reverse otherwise. This is a permanent feature of a share tenant family in the sense that this decision stands irrespective of whether it is a pure tenant or a predominantly wage earner or a combination of both. These results go in conformity with those of Pant (1983) when the landless household is either a pure tenant or a predominantly wage earner. But in the case of mixed status of the household i.e., when it is simultaneously a wage earner as also a share tenant, Pant (1983) had the result that a change either in the size or in the composition of the landless household will have no effect on its lease-in demand. In support of this result it had to asume that labour market is perfect and there is perfect substitutability between dependent labour and wage employable labour within the family as far as division of responsibilities between different cultivating activities are concerned. This is a questionable assumption indeed. For a landless household deriving income from both wage market and tenant cultivation, it seems more rational to make a balance between these two sources of income in the sense that wage income is not affected, irrespective of what the wage rate is, by tenancy cultivation. For such a household, if tenancy cultivation is organised primarily with the objective of providing income earning opportunities to other members of the family who are otherwise disqualified in the wage market, wage income is better to be treated as an income in addition to tenancy income. As a result, it is reasonable to believe that tenancy should not be allowed to disturb wage income as far as possible. Accordingly, perect substitution between dependent labour and wage employable labour within the landless household (Pant,1983) is not desirable. In fact, current results go in perfect conformity with those of Pant (1980). The trouble with the results of Pant (1983) is possibly with the fact that it wants to examine the effect of change in the size and composition of the landless household on its leased-in demand in the light of leasing-out tendency of the landowning household i.e., leasing-out activity and the leasing-in tendency of these two households are taken together.

Subsequently the current exercise has demonstrated that lease-in demand from the landless household will go down with every increase in wage rate as also with every increase in land preparation cost. But as expected, a landless household with more managerial efficiancy will want to have more leased-in land; this has perfect harmony with Pant (1983). But regarding the other two results i.e., with reference to effect of change is wage rate and change in land preparation cost on lease-in demand, Pant (1983) had the opposite result; again this is possibly due to the effect of joint consideration of lease-out and lease-in activities together.

Since the prime focus of the paper was to examine the effect of change in the family resources, particularly those varieties which are either not marketed or marketed under imperfect conditions, on the leased-in demand for the landless household, the behaviour of the landowning household has received less attention. It is the land market where leasing activities of these two households directly interact with each other through the variable rental rate. When the lease-in demand of all the landless households get changed due to change in composition of family resources, it will have its effect on rental rate, which in turn will affect the lease-out behaviour of the landowning households. There is another market–wage market– where these two households do interact with each other. Again these two varieties of interactions between these two sets of families are not independent; it will affect demand for labour in the wage market and as a consequence wage rate will be affected. Accordingly wage income of the landless households will get disturbed.

As far as managerial efficiency of the landless households is concerned, the higher the efficiency, the more is the demand for lease-in land and accordingly less is the supply of wage labour. That means efficient households will prefer to be pure tenants. Again by moving one step further, it can be said that efficient pure tenants will prefer to be fixed rent tenants rather than share tenants. Thus change in the managerial efficiency has a far reaching consequence as far as the landless household is concerned. These are some of the issues on which extension of the current model is under preparation.

NOTES

[1] For justification of this assumption refer to footnote 2 of Pant (1983)

[2] Among all the strenuous activities, only land preparation has been specified since it can vary from high rainfall (fertile) areas to low rainfall (less fertile) areas.

[3] At times it may appear that a better way to bring in m is to relate it with l_2 instead of $\bar{l}$, particularly so, since both $\bar{l}$ & l_2 perform only routine cultivating activity. But since the difference $l_2 - \bar{l}$ is the wage employable labour performing routine activity, it will be wise for any household to get as much ordinary work performed by the dependent labour as possible, the more it is so, the less is the requirement of wage employable labour for routine

activity. As a result, it appears more proper to define managerial capability in terms of $\bar{l}$ instead of l_2

[4] For simplicity it may be viewed as every fall in $K \Longrightarrow$ increase in E where E stands for efficiency.

[5] Second order conditions have been verified.

[6] Second order conditions have been verified

[7] Since the current exercise is primarily concerned with effect of change in the size and composition of the household labour force on leased-in demand, only this result has been reported. Similar results with respect to other attributes could have been derived as in the previous case [refer to results (4a) to (4d)].

[8] Second order conditions verified

[9] Similar results with other attributes are obtained but not reported here.

[10] It may be worthwhile to reconcile that in the definition of K, labour hours deserving managerial supervision was mentioned as $\bar{l}$ and not l_2 where $\bar{l}$ was the dependent labour, l_2 was the labour engaged in routine operation and $l_2 - \bar{l}$ was the wage employable labour engaged in routine activity.

[11] Between the production functions F and G, except the attributes there is no fundamental difference.

[12] Second order conditions have been verified.

[13] By definition $l_3 = \bar{\bar{l}} - l_2 + \bar{l} - ah - (Kh/\bar{l})$ assuming demand-supply equality in the labour market

REFERENCES

Bardhan, P.K., 1976, Variation in extents and forms of agricultural tenancy An analysis of Indian data across regions and over time, *Economic and Political Weekly* XI, Sept., 37-38

Bell, Clive, 1976, Production conditions, innovations and the choice of lease in agriculture, *Sankhyā*, C, Dec.

Bliss, C. J. and N. H. Stern, 1982, Palanpur : Studies in the economy of a North Indian village, Oxford University Press.

Pant, Chandrashekar, 1980, Contractual Agreements in agriculture : Some theory and empirical evidence, Unpublished Ph. D. thesis (Indian Statistical Institute, Calcutta)

Pant, Chandrashekar, 1983, Tenancy and family resources, *Journal of Development Economics*, 12, 27-39.

Layoff Compensation and Entry Deterrence

Bibhas Saha & Anindya Sen

1. Introduction

An incumbent firm can deter the entry of new firms threatening to produce more output in the event of entry (i.e. limit pricing) only if such a threat is credible. The literature on entry deterrence has shown that there are various ways limit pricing can be made a credible threat, e.g. by choosing excess capacity [Spence (1977). Dixit (1981), Spulber (1981)], through high exit costs [Eaton and Lipsey (1980)], etc. For a survey of this literature see Gilbert (1989). The entry deterring strategies often take the form of some irreversible expenditures made by the incumbent at the pre-entry stage so that his costs are reduced in the post-entry situation or his determination to fight becomes stronger.

An alternative way of deterring entry is through self-enforcing contracts which was suggested by Aghion and Bolton (1987). In their model an incumbent monopolist seller writes a contract with his buyer, who is a third party in the entry game, specifying two prices: one for buying from him and the other for not buying from him. The second price is essentially a penalty for switching to the entrant. This arrangement is shown to deter entry. Dewatripont (1989) uses a similar approach by introducing a labor union as the third party. The incumbent and the labor union agree to a renegotiation-proof wage and employment contract in the presence of asymmetric information.[1] Incentive compatibility constraints require specifying output at a high level and this may deter entry.

In this model, we explore in detail the possibility of entry deterrence by means of layoff compensation (or severance pay).[2] Layoff compensation is introduced as part of a labour contract offered by the incumbent to its workforce which is presumably unionized. According to this contract holding excess workers or not producing output at a precommitted level is costly.[3] Limit pricing may then become a less painful alternative than entry accommodation. Therefore, by choosing such contracts the incumbent can pose a credible threat to the entrant. However, whether entry can be deterred or not would depend largely on the relative cost parameters of the rival firms. In essence, the provision of layoff compensation creates barriers to entry by creating barriers to exit for the incumbent firms.

It should be noted that layoff compensation can be a part of labour contract not only when the firm volunteers to pay, but also when the workers are unionized and demand such compensations. Layoff compensation can also be committed to non-unionized (or spot market) workers. But as Dewartnpont (1987) argues, using contracts with unions has some advantages. Union contracts are more publicized than non-union contracts and hence, observable to potential entrants. Also it is true that the existence of union is a credible guarantee against unilateral violation of the contracts.

Similar credibility can also be achieved when labour legislations require firms to pay compensation. This possibility is particularly interesting in the context of India labour regulations. In india, the Industrial Employment Act, 1946 specifies that a firm (in the organized sector) having a workforce of 50 regular workers, will have to pay a severance pay of '15 days' salary for every year worked' to its retrenched workers. Another act curtails the firm's right to retrench for all firms having a workforce of 100 or more workers. Similar regulations are also common to many European countries.

Our result will be useful to understand to what extent a layoff compensation regulation can reduce the competitiveness of a market by creating barrier to entry.

The key to our result is the way the cost structure of the incumbent firm is altered by layoff compensation. Since a worker, once hired, cannot be costlessly retrenched, the firm perceives a fixed cost on its entire workforce, by the amount of layoff compensation. This may render the duopoly outcome (in which case it has to retrench workers) more unprofitable than producing the limit output. We found that under some reasonable restrictions on wage and severance pay, entry deterrence is indeed the unique outcome. Even the well-known case of the incumbent and entrant having identical marginal costs which usually gives entry accommodation, gives rise to the equilibrium of entry deterrence.

However, to create a credible threat against entry, the incumbent may need to choose a large workforce, but in the entry deterrence equilibrium he might have to retrench some workers when the entrant stays out. Therefore, it has to be examined whether it is optimal for the incumbent to hire excess workers. Interestingly, it will be optimal if the rate of layoff compensation is below a critical level.[4]

The paper is organized as follows. Section 2 presents the model. In Section 3, the no commitment case is discussed. Section 4 analyzes the role of layoff compensation in entry deterrence. The choice of workforce and equilibrium for the entire game are analyzed in Section 5. Additional comments are provided in the concluding section.

2. The Model

There are two firms - the incumbent, denoted by subscript 1 and the entrant, denoted by subscript 2. The incumbent was producing the monopoly output by employing an optimal workforce. In the face of an entry threat, he may be interested in altering his behavior. The incumbent's behaviour and subsequent developments are captured in a three-stage entry game.

The first stage of the entry game concerns the incumbent's *adjustment* of workforce and *choice* of severance pay. Choosing a *positive* severance pay will be considered as making a commitment. One underlying assumption is that previously there was no severance pay commitment or requirement on the part of the firm.[5]

Since the incumbent already had a workforce, any adjustment in the number of workers can be understood as a renewal or a revision of contract. Suppose the firm wants to hire a workforce of size N which may be bigger or smaller than its current workforce. Then N workers are chosen by the incumbent from a union size of M identiacal workers[6] It can be seen that there are natural bounds on N . N cannot be less than the duopoly level and will not exceed the 'limit output' level.

If a member of the union does not fall in the group of N, he works in the spot market and gets a utility v. On the other hand, if he is included in N, then he waits for another draw. The firm may or may not use (or *employ*) its entire workforce. Out of N, only L workers are employed ($L \leq N$) and the rest are paid layoff compensation. Note that we are making a distinction between *workforce* and *employment* within the firm.

Each worker is assumed to work for 1 manday. For simplicity, w is fixed from outside. Thus a worker's expected utility is

$$Eu = \{(L/N)\ u(w,1) + (1 - L/N)\ u(t,\theta)\}\ (N/M) + [1 - (N/M)]v$$

Assume $v \leq u\ (t,\theta) \leq u\ (w,1)$ and $w > t$ and $u(.,.)$ increasing and concave in w. The assumption that $u(t,\theta) \leq u\ (w,1)$ can be guranteed and is needed for the firm to eliminate the undesirable possibility of 'involuntary employment'. But $v \leq u\ (t,\theta)$ may not be ensured. As long as $u(w,1) \geq v$, there may be an incentive for the worker to stay in the union. If v increases beyond $u(t,\theta)$, the outcome of the model may change[7] So we make a simplifying assumption that v is sufficiently small. Finally, it is important that the contracts are binding and are publicly observable.

In the second stage, the entrant decides whether to enter or to stay out. If he enters, he has to incur a sunk entry cost F which is assumed to be very small, close to zero. His objective function is

$$\pi_2 = Py - cy - F \qquad (1)$$

where c=MC=AC and y refers to his output. By staying out he ensures no-profit-no-loss.

In the third and final stage, if the entrant has entered the incumbent decides whether to 'fight' by producing the limit output or to 'share' the market. From the first strategy, his payoffs are π_1^{*} in the case of commitment and π_1^{*n} in the case of no commitment (all variables with superscript n refer to the no commitment case). By playing the second strategy--'share'-- the two firms will be producing the Cournot-Nash outputs. In this case, the incumbent receives profits π_1^{d} and π_1^{dn} , respectively. But if the entrant has chosen to stay out, then the incumbent enjoys his monopoly resulting in profits π_1^{m} and π_1^{mn} respectively. Since commitment is costly, the incumbent's profits are related in the following manner:

$$\pi_1^{d} < \pi_1^{dn}, \qquad \pi_1^{m} < \pi_1^{mn}$$

but $\pi_1^{*} = \pi 1^{*n}$ as at the limit output all workers will be fully employed and the layoff cost would be zero. Therefore, regardless of commitment, profit in this case will be same.

The incumbent's technology is assumed to exhibit CRS with labor being the only input for production. Let the production function be x=AL. Normalizing A=1, we can rewrite x=L. In the commitment case, the incumbent's problem may be stated as

$$\text{Max } \pi_1 = Px - mx - tN \tag{2}$$

where m = w-t.

In the no commitment case, the incumbent's profit is $\pi_1 = p_x - w_x$. The market demand curve is assumed to be linear : p = a – (x+y). It is assumed that a>c,w. We want to see under what conditions the following strategy combition forms a subgame perfect Nash equilibrium : commitment, stay out and fight.

3. The No-commitment Case

Let us first consider the no commitment case where the layoff cost is zero, i.e. t=0. In this case, there is no real distinction between workforce and employment. A worker remaining in the workforce and being unemployed, does not receive any layoff compensation. So he would prefer to work outside in the event of unemployment. Therefore, the firm does not choose a workforce bigger than its employment size. This case would be used as a benchmark to evaluate the benefits of commitment.

In order to ensure subgame perfectness, we proceed backward by determining the last mover's optimal strategy first. In the last stage of the game, the incumbent chooses whether to fight (i.e. to engage in limit pricing) or to share (i.e. to produce the Cournot output). The reaction curves of the incumbent and the entrant are, respectively,

$$x = (a-w-y)/2 \quad \text{and} \quad y = (a-c-x)/2. \tag{3}$$

If the incumbent plays 'share', the resulting duopoly outputs and profits are

$$x^{dn} = (a+c-2w)/3, \quad y^{dn} = (a+w-2c)/3 \text{ and}$$

$$\pi_1^{dn} = (a+c-2w)^2/9, \quad \pi_2^{dn} = (a+w-2c)^2/9. \tag{4}$$

We assume

$$a+c-2w \geq 0, \quad a+w-2c \geq 0 \tag{5}$$

F is assumed to be small enough to ensure positive profit for the entrant in a market sharing equilibrium.

On the other hand, if the incumbent plays 'fight', the resulting outputs and profits are $X^{*n} = a-c$, $y^* = 0$ and

$$\pi_I^{*n} = (c-w)(a-c), \qquad \pi_2^* = -F \tag{6}$$

which indicates that if the 'fight' strategy is credible, the entrant will prefer to stay away. To check the credibility of the fight strategy, we have to examine whether $\pi_I^{dn} < \pi_I^{*n}$. As long as $0 \leq w < (a+c)/2$, $\pi_I^{du} > 0$ and for $\pi_I^{*a} > 0$, we must have $w < c$. So the credibility of the fight strategy has to be examined only for $w < c$. Let us define

$$g(w) = \pi_I^{dn} - \pi_I^{*n}$$

as a function of w parameterized by c and a. Substituting from (4) and (6) we obtain

$$g(w) = \{(a+c-2w)^2 - 9(c-w)\ (a-c)\}/9 \tag{7}$$

which is quadratic in w. After rearranging terms,

$$g(w) = (4/9)\ \{w-(2c-a)]\ [w-\{(5c-a)/4\}] \tag{7}$$

The two roots of g(.) =0 are given as

$$w^+(c) = (5c-a)/4 \text{ and } w^-(c) = 2c-a \tag{8}$$

Clearly, $w^-(c) < w^+(c) < c$ Since $a > c$. The function g(w) attains negative values for all w lying in the interval $[w^-(c), w^+(c)]$ implying $\pi_I^{dn} < \pi_I^{*n}$.

So for entry deterrence, w has to be sufficiently less than c. If w is not the above range, the incumbent would prefer to share. But note that the critical range of w depends on the entrant's cost parameter c. Therefore, the comparative costs of the firms are important in determining the optimality of limit pricing as shown in Table 1.[8]

w/c	$0 \leq c \leq a/5$	$a/5 < c$
$\max\left[w^-(c),\ 0\right] \leq w \leq w^+(c)$	Entry Accommodated	Entry Deterred
$w^+(c) < w \leq (a+c)/2$	Entry Accommodated	Entry Accommodated

Table 1.

Case 1. Low cost entrant: In column 1 of Table 1, we identify the case where entry cannot be deterred. If the entrant's MC is less than or equal to a/5, the incumbent, regardless of his cost conditions, will find sharing as the opimal strategy. This is evident from the fact that $w^{+}(c) < \phi$ for $c < a/5$. At all positive w, $\pi_1^{dn} < \pi_1^{*n}$. Thus a/5, can be considered as a critical value of the entrant's MC and an entrant with MC below (above) this level will be referred to as a low (high) entrant.

Case 2. High cost entrant: On the other hand if the entrant's MC exceeeds the critical value a/5, the incimbent can deter entry provided his MC (i.e.w.) is within the range $[w^{-}(c), w^{+}(c)]$. This shows that the incumbent will find it optimal to deter entry if the entrant suffers an *absolute* cost disadvantage (c being greater than a/5) and if the incumbent enjoys a relative cost advantage $w^{+}(c)$ being stricly less than c). The second column of Table 1 specifies the ranges of w within which the incumbent will deter or accommodate entry.

4. The Case of Commitment.

Now let us consider the commitment case where the incumbent chooses a workforce with a positive layoff cost i.e. $t > 0$. Since we want to see whether commitment enables the incumbent to produce the limit output x^* in a credible way, let the firm set N equal to x^*, i.e. $N = (a-c)$.

Assumption : N is set at the limit output level, $N = a-c$.

We are interested in analyzing whether an imposition of positive layoff cost along with a precommited workforce alters any outcome on Table 1.

Again, starting backward from the last stage, we first solve for the incumbent's optimal strategy. The duopoly outputs and profits are

$$x^d = (a+c-2m)/3,\ y^d = (a+m-2c)/3 \text{ and}$$
$$\pi_1^d = \{(a+c-2m)^2/9\} - tN,\ \pi_2^d = \{(a+m-2c)^2\} - F, \tag{4'}$$

For model consistency we set

$$a+c-2m \geq 0,\ a+m-2c \geq 0. \tag{5'}$$

Note that the duopoly output and profit expressions in the present case differ from the earlier case only to the extent that w in the earlier case is replaced by m ($\equiv w-t$) in the present case. But the incumbent's limit output and profit will remain unchanged as given in (6), assuming $N = a-c$.

The introduction of layoff compensation has two effects: (i) it reduces marginal cost from w to m given a cost advantage to the incumbent in the post–commitment stage; (ii) but it also adds a fixed cost tN. Each worker has to be paid at least t regardless of employment. But the additional wage cost per employed worker (which is the only variable cost in this model) is just the difference between wage rate and the rate of layoff compensation. The fixed cost of layoff compensation is unavoidable even when the firm closes down.

This is represented in Figure 1. While a reduction in marginal cost shifts out the reaction curve of the incumbent from RR to R'R', the imposition of fixed cost, interestingly, does not truncate it (or introduce discontinuity). The reason is that the relevant minimum profit level is no longer zero, but –tN. The firm has to incur bigger loss than –tN before it exits. It can be shown that its profit is greater than '–tN' everywhere on its reaction curve.[8]

The new duopoly output point moves from point D to D'. In order to justify his strategic choice of t, we need to consider his profits only at the two output levels x* (the limit output) and x corresponding to D. While profit at x* is unchanged, it may increase or decrease at D'. So it is possible that point D' might be positioned above the isoprofit curve passing through x* indicating that duopoly profit is less than the 'limit profit'.

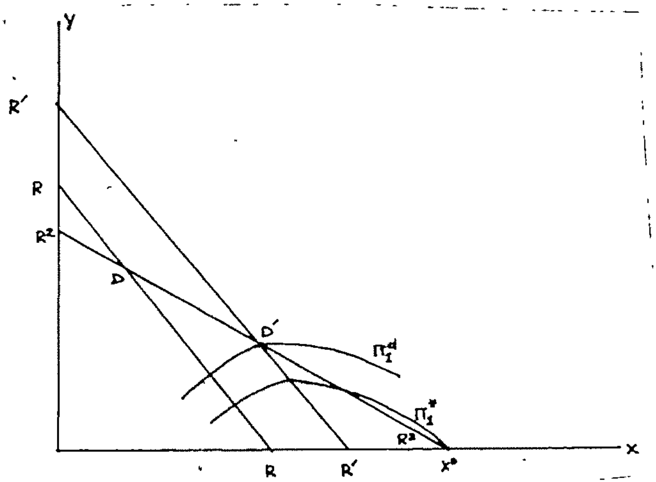


Figure 1

Formally, redefine $g(m) = \pi_1^d - \pi_1^*$ as before and obtain

$$g(m) = (4/9)[m-(2c-a)][m-\{(5c-a)/4\}] \qquad (7'')$$

Note that (7") and (7') are different only to the extent that w in (7') is replaced by m in (7") and $0 \leq m \leq w$. Therefore, the two roots of g (m,c) = 0 would be similar to $w^+(c)$ and $w^-(c)$:

$$m^+(c) = (5c-a)/4, \quad m^-(c) = 2c-a. \qquad (8')$$

Furthermore, g (m) < 0 for m lying in [$m^-(c)$ and $m^+(c)$] ; otherwise g (m) >0.

Effectiveness of layoff compensation : As long as the 'net wage rate' m is in the critical range given by (8'), the incumbent will find it optimal to produce the limit output. Note that the critical interval of m is identical to that of w given in (8); but they correspond to different wage rates. Although the incumbent can influence the value of m by suitably choosing t, he is constrained by the level of w. We discuss the implication of the critical interval of m in two subcases.

(a) *The case of* $w \geq w^+(c)$: The most interesting result in this section is that even if $w > w^+(c)$, the incumbent will have incentives to deter entry by choosing a t from the interval $[w - w^+(c), w - w^-(c)]$. The lower bound ensures that $m < m^+(c)$ and the upper bound $m > m^-(c)$ which in turn guarantees that limit pricing will be optimal. This result holds even if w > c. Since t increases firm's costs, the incumbent choose $t = w - m^+(c)$.

In addition to g (m) < 0, we must check whether the firm would ex post prefer 'fight' to 'exit'. By exiting it loses t (a-c) and by fighting it earns (a-c)(c-w) if c>w and loses (w-c)(a-c) if w>c. When w>c, its loss from "fight' will exceed its loss from 'exit' if w-t > c. But this is not possible as $t \geq w - w^+(c)$ and $w^+(c) < c$.

Therefore, having made a commitment, the incumbent can deter entry even if he suffers a cost *disadvantage* vis–a–vis the entrant. This result is markedly different from the no–commitment case, where the incumbent must have a relative cost advantage over the entrant to carry out limit pricing.

(b) *The case of* $w < w^{+}(c)$: There are two sub–cases :

$w < w^{-}(c)$ and $w > w^{-}(c)$. First, let us consider $w < w^{-}(c)$. In this range the incumbent would find commitment to be useless, because $m < m^{-}(c)$ and market sharing rather than deterring entry becomes optimal. In the case of $w > w^{-}(c)$ entry deterrence is possible without commitment. So in both cases, t will be set at zero.

Note that if the rate of layoff compensation is institutionally fixed, the entry outcome will depend on the level of t as well as on the levels of w and c. For instance, if t was greater than $\left[w - w^{-}(c)\right]$, then the incumbent cannot deter entry even though he is perfectly capable of doing so in an unconstrained environment when $w > w^{-}(c)$.

Once again, the attractiveness of the limit pricing strategy depends on the entrant's costs. If the entrant's MC is lower than the critical level a/5, the incumbent would not be interested in deterring entry. This is obvious from the fact that $m^{+}(c)<0$ for $c<a/5$. Since we have restricted t to be less than w on the ground that an employed worker is generally better off than a laid off worker, m cannot be negative in our model. In other words, the low cost entrant who was not deterred in the no commitment case, will still remain undeterred and the incumbent's commitment will be of no use. Deterring entry is optimal only when the entrant has high cost, i.e. $c>a/5$.

Thus it can be said that layoff compensation is useful in some but not all cases. Recalling Table 1, outcomes specified in column 1 of Table 1 will remain unchanged also in the commitment case. Only the outcome entered in the second row of the second column of can be changed choosing an appropriate rate of layoff compensation. This signifies the scope of entry deterrence through layoff compensation.

Proposition 1:

Let the entrant have marginal costs $c>a/5$. Then there is a unique subgame perfect Nash eqilibrium in which the incumbent chooses $t=w-w^+(c)$ [or slightly greater than $w-w^+(c)$] and the entrant does not enter.

5. Optimality of commitment

The proposition 1 from the previous section, describes an outcome where the incumbent will be strategically holding excess workers. The size of the hired workforce is N(=a-c), but the actual level of employment, after the entrant stays out, will be at the monopoly level, $(a-m)/2 = \{a-w^+(c)\}/2 = 5(a-c)/8$. The number of unemployed workers will be $3(a-c)/8$. This raises a question whether making such a costly commitment is optimal.

Usually, the benefits of commitment (and subsequent entry deterrence) are considered to be large if the threat of entry is one shot or going to be infrequent in the long run. But if entry threat is persistent, then benefits are likely to be outweighed by costs of commitments.

To get a clear answer, we have to compare the incumbent's monopoly profit that is obtainable only after deterring entry vis-a-vis his duopoly profit when no prior commitment is made i.e. $\pi_I{}^m = \{(a-m)^2/4\}-tN$ with $\pi_I{}^{dn}=\{(a+c-2w)^2/9$. We find that there is a critical level of the rate of layoff compensation above which commitment would appear as too costly.

For successful entry deterrence m cannot be more than $m^+(c)$ which corresponds to the lowest required layoff compensation rate $t^*=w-w^+(c)$. So allowing $t=t^*$ and setting $N=(a-c)$, we obtain

$$\pi_I{}^m = [\{(a-w^+(c)\}^2/4]-\{w-w^+(c)\}\ (a-c) \tag{9}$$

Define $f(w) = \pi_I{}^{dn}-\pi_I{}^m$

$$= [(a+c-2w)^2/9] - [\{a-w^+(c)\}^2/4] + [\{w-w^+(c)\}\ (a-c)] \tag{10}$$

The two roots of $f(w)=0$ are approximately equal to $(a+15c)/16$ and $(37c-21a)/16$[10] The second root is less than $w^-(c)$ and therefore can be ignored. Let us denote the other root as $w^*(c)$:

$$w^*(c) \simeq (a/16)+(15c/16) \tag{11}$$

which is clearly greater than c. Note that $f(w)>0$ for $w>w^*(c)$.

Nash equilibrium : The above discussion enables us to characterize the Nash equilibrium of the game. Whether the incumbent will decide to commit to layoff compensations or not, will depend on the optimality of limit pricing that would be carried out subsequently if the entrant enters. As we have seen in the previous sections, limit pricing is never optimal (regardless of commitment) if the entrant's MC is sufficiently low, i.e. $c < a/5$. Therefore the incumbent would not make any commitment and will not hire a workforce above the duopoly level and subsequently would accommodate entry. But when $c > a/5$, there are prospects of deterring entry with or without commitment depending on the incumbent's cost level. If w is in the interval $[w^-(c)w^+(c)]$, then he can deter entry without commitment.

Entry deterrence under commitment : The more interesting case arises when the entrant has high cost (i.e. $c > a/5$) and the incumbent *does not have* a relative cost advantage [i.e. $w > w^+(c)$]. We know that there exists a positive layoff compensation rate (which must also exceed $[w-w^+(c)]$ that would enable the incumbent to deter entry. Since commitment is costly, such strategy is unlikely to be optimal at all levels of the wage rate. In fact the commitment strategy is optimal for the incumbent as long as his wage cost does not exceed a critical level $w^*(c)$ as given in (11). An incumbent who faces a wage cost $w < w^*(c)$ will find it optimal to deter entry by choosing a large workforce and a rate of layoff compensation $t = w-w^+(c)$. Clearly for an incumbent with $w > w^*(c)$ making such commitments would be too costly.

Therefore, for all w satisfying $w^+(c) < w \leq w^*(c)$ and $N = a-c$, the unique subgame perfect Nash equilibrium involves commitment of $w-w^+(c) \leq t$ on the part of the incumbent and not entering the market on the part of the entrant as he finds limit pricing a credible threat. This case is presented in the second column of Table 2.

For any w exceeding $w^*(c)$ or any t not satisfying the above restriction, the Nash equilibrium involves entry and no commitment on the part of the incumbent.

Table 2

	$0 \le c \le a/5$	$a/5 < c < a$
$\max[0, w^-(c)] < w \le w^+(c)$	Entry Accommodated	Entry Deterred
$w^+(c) < w \le w^*(c)$	Entry Accommodated	Entry Deterred under Commitment
$w > w^*(c)$	Entry Accommodated	Entry Accommodated

Proposition 2

Let $c > a/5$. If $w^+(c) < w \le w^*(c)$, then the incumbent will make a commitment by choosing $t = w - w^+(c)$ and subsequently entry is deterred. If $w \ge w^*(c)$, t will be set equal to zero and entry will occur.

6. Conclusion

We have analyzed the possibility of deterring, entry by means of self-enforcing contracts between the labor union and the incumbent. Commitment takes the form of layoff compensation paid to the excess workers held by the incumbent. We found that while layoff compensation is useful in deterring entry over a significant range of parameters, the rate of layoff compensation has to be restricted to make this commitment optimal. Thus, even in a very simple model, we have been able to identify the incumbent's incentive to prefer a unionized workforce in some cases.

Our model adopted a principal agent framework in modeling the firm-union relationship. However, both the wage rate and the layoff compensation rate may be determined by bargaining between the union and the management. The management would bargain for lower wage, but from the point of view of entry deterrence, it may be interested in higher layoff compensation rate.

Again, the paper assumed constant returns to scale. It might be interesting to analyze situations of increasing returns to scale and whether the provision of layoff compensation serves to reinforce a situation technologically favorable for entry deterrence.

NOTES

1 The impact of labor union on the incumbent's incentive to deter entry is also discussed in Dewatripont (1987) where he suggests that 'high' severance pay as part of employment contract can facilitate entry deterrence

2. Prevalence of layoff compensation, also known as supplemental unemployment benefit, has been cited in the labor literature For example, Oswald (1986) reports that throughout the seventies, approximately half the workers in the US manufacturing industry had supplemental unemployment benefit. The 1982 contract signed by the UAW and General Motors had provision of layoff compensation as high as approximately 95 percent of the after-tax pay of an employed worker. Traditionally, the US auto industry has provided quite a high rate of layoff compensation - 70 percent of regular pay.

3 Note that in our model "layoff compensation" is more appropriate than "severance pay" Commitment entails holding excess workers and therefore a positive payment made to them should be called layoff compensation Severance pay in our model will be zero.

4 However, it can be shown that entry deterrence does not necessarily require holding excess workers. There are parameter configurations when the entrant is deterred by the incumbent who chooses a workforce between his 'pre-entry' monopoly level and the 'limit pricing' level provided that the entrant has a moderate fixed cost and this workforce will be used fully in the entry deterrence equilibrium.

5. If however, the firm is required by regulation to pay a minimum severance pay, commitment will have to understood with resptect to a rate of severance pay that is above the statutory minimum.

6. Perhaps it is more realistic to think that the firm randomly selects only the new workers and retains the old ones, when it wants to increase the workforce However, since all workers are identical, the random selection rule does not affect the union's or firm's objectives. So this rule is used for model simplication.

7. If $v > u(t, 0)$ and v is sufficiently low, then all our results hold. But if v is sufficiently large, then L/N also has to be sufficiently large to satisfy the union's individual rationality constraint. In that case, some of our results may change For example, if $v = u(w, 1)$ as an extreme case, then it is optimal for the firm to choose union and set output at the Stackelberg level and allow entry See Saha and Sen (1990).

8 Note that in Table 1 we have not considered w less than $w^{-}(c)$. The reason is that $w^{-}(c) \leq 0$ for $c \leq a/2$. For $c > a/2$, though $w^{-}(c) > 0$, $w^{-}(c)$ violates one of the constraints in (5)

9. The profit of firm 1 along its reaction curve is

$$\pi_1 = (a-w-t-y)\,(a-w+t-y)\,/4 - t\,[N - (a-w+t-y)\,/2].$$

Now let $y = \alpha\,(a-w+t)$ where $0 < \alpha \leq 1$ so that by changing α we can identify the level of y At $\alpha = 0$, $y = 0$ and π_1 corresponds to the monopoly profit. At $\alpha = 1$, y corresponds to firm 2's limit output. Then it is easy to check that π_1 is declining in α and at $\alpha = 1$, $\pi_1 = -tN$. Therefore, if firm 1 makes losses anywhere on its reaction curve, the loss would be less than '$-tN$'.

10 The exact values for the two roots are

$$w^* = [\{(-10a+26c) \pm (a-c)\,(117^{1/2})\}/16].$$

We obtain the approximate value of w^* as shown in the text by setting $117^{1/2} \simeq 11$ to get less cumbersome expressions. It can be easily checked that this makes no difference to the results.

REFERENCES

Aghion, P., and Bolton, P (1987): "Contracts as a Barrier to Entry" American Economic Review 77 : 388–401.

Dewatripont, M. (1989): "Commitment through Renegotiation-proof Contracts with Third Parties". Review of Economic Studies 55: 377–390.

-----------(1987); "Entry Deterrence under Trade Unions". European Economic Review, 31: 149–156.

Dixit, A. (1981): "The Role of Investment in Entry Deterrence". Economic Journal 90: 95–106.

Eaton, B , and Lipsey, R (1980): "Exit Barriers are Entry Barriers : The Durability of Capital as a Barrier to Entry". Bell Journal of Economics 11: 721–729.

Gilbert, R. (1989): "Mobility barriers and the value of incumbency" in R Schmalensee and R. Willing (eds): *Handbook of Industrial Organization*, (North–Holland, Amsterdam).

Oswald, A. J. (1986): "Unemployment Insurance and Labor Contracts under Asymmetric Information : Theory and Facts". American Economic Review 76: 365–377.

Saha, B., and Sen, A. (1990): "Layoff Compensation asCommitment for Entry Deterrence". Discussion paper no. 39, Indira Gandhi Institute of Development Research.

Spence, A. M. (1977): Entry, Capacity, Investment and Oligopolistic Pricing". Bell Journal of Economics 8 534–544.

Spulber, D. F. (1981): "Capacity, Output and Sequential Entry". American Economic Review 75: 897–899.

Non-Walrasian Static Exchange Problem

Ishita Mukhopadhyay

The paper considers an illustration from an exchange problem in a non-Walrasian reference frame, where compatibility between the applied mathematical approach and the implied economic logic is explored. We will refer to implied economic logic as the "subjective construct" of the problem. It is this "subjective construct" which is acted upon by mathematical approaches. The implication of the compatibility test is the fact that the in-built structure of a particular economic problem allows certain mathematical constructs and not all. The specific construct considered here has been acted upon by one mathematical approach. It may happen that a single "sujective construct" is acted upon by more than one mathematical approaches. Compatibility tests will obviously negate some approaches and may not even retain a single one of the existing approaches. Thus even coexistence of a single mathematical approach and a single subjective construct does not alone prove compatibility between the two. Hence the compatibility exercise performed in this paper is justified. Section 1 describes the static non-Walrasian exchange problem. Section 2 outlines the subjective construct derived from the description. Section 3 discusses the applied mathematical method with the intuitions embodied in it. Section 4 performs the direct compatibility exercise. Section 5 attempts to provide an alternative mathematical approach satisfying the compatibility tests.

1. Description of the Problem

A standard non-Walrasian exchange problem starts with a macro aggregative mismatch between demand and supply. This is analysed by seeking a micro foundation in terms of basic units of production and consumption. The entire analysis starts off from the macroeconomic problem as an observed economic phenomenon. Starting point is a non-matching exchange. The endpoint of the analysis is an observed economic solution. The solution is a trade of mutual satiation.

The path from the macroaggregative mismatch to a trade of mutual satiation is as follows. The economic agents i.e. the sellers and buyers of different commodities participate in a demand-supply contract. The contract allows the agents to reach a state of mutual satiation maintaining the macroaggregate gap. As this is a direct counter example to Walrasian state, the need for a non-Walrasian analytic framework becomes evident. Clower (1965), Barrow and Grossman (1976) came up with non-Walrasian presentations. They were followed by Benassy (1978), Dreze (1975), Malinvaud (1977), Grandmont (1982), Grandmont and Laroque (1976).

Given Walrasian tatonnement process, it is not possible to conceive of a mismatch between aggregate demand and aggregate supply. Hence the earliest attempts by Hahn and Negishi (1962) and Uzawa (1962) conceptualised the situation in terms of non-market clearing fixed prices. Clower (1965) looked at the impact on optimising behavior of agents if there exists such prices. This non-market clearing fixed prices become associated with effective demands. Walrasian situation implies notional demands, whereas non-Walrasian situation brings in effective demands. Clower (1965) pointed out that in such situations of rigid prices, adjustments take place through quantities. This rise to a series of macroeconomic models assuming fixed price structures with quantitative adjustments.

A second variant of such a system is based on conceptualisation of identical individuals. The consumers or buyers have identical preference patterns and endowments. The producers or sellers face identical technology. A monetary economy is assumed where money is a numeraire, a medium of exchange, and a store of value. Trade of a single commodity between two agents – one buyer and one seller is defined as a partial trade. Partial trade balance is an aggregative counterpart of partial trades. Incorporation of money into individual endowments of buyers is not necessary to get the results as shown by Benassy (1978). So the present formalisation excludes monetary intrusion.

The formalisation used is a structure of many commodities and many individuals. An economy E(n, m) is defined by n goods and m individuals. The goods include money as the n-th good, and (n-1) goods are traded against the n-th one.

An individual "i" is either on the short side or on the long side of the market for commodity "j". The problem arises due to (a) spilling over of information from market for one good to market for the other good and (b) interrelated demands and supplies of the commodities in the various markets. As a clarification, one may consider the situation when an individual being an excess demander of a particular commodity is affecting his demand for another commodity. This brings in the idea of a quantitative constraint as felt in the different markets by an individual. The constraints are defined as perceived constraints. These perceived constraints serve to derive effective demands from notional demands.

The process of reaching a trade of mutual satiation is a simultaneous one. A constrained perception rule derives perceived constrained sequences corresponding to different individuals and different commodities. These perceived constraints determine the effective demands. It is the actual trades expressed via this mechanisms which determine the perceived constraint sequence by the same constraint perception rule. The situation is illustrated in Fig. 1.

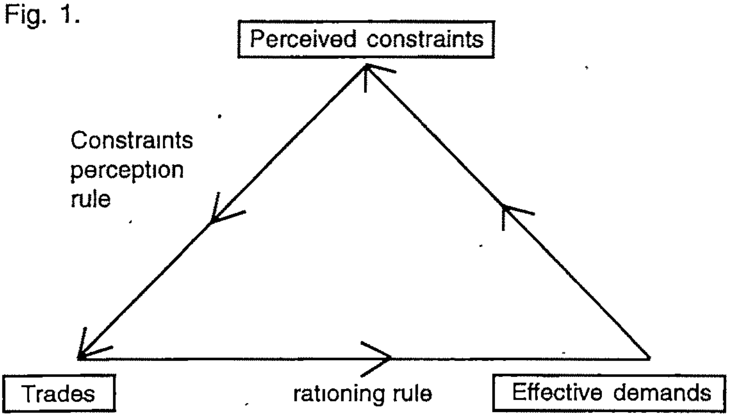


Fig. 1

The process is continued until we have a common convergence of the sequences. Modifications and further modifications on the effective demand sequence then stop at the common point of convergence. The entire analysis of existence of this common point, which has been termed as the equilibrium point by Benassy (1978), its uniqueness, stability, comparative statics then follows in.

2. The Subjective Construct Within

The subjective construct within the description presented in Section 1 is now derived. The description is a micro-based analysis of a macroeconomic problem. The problem is observation of the fact that the macroaggregate demand is not matching with supply. The phenomenon is an observable in a macro-level. A simultaneous observable phenomenon is that the economic agents participating in micro-level demand-supply contracts i.e. the sellers and buyers reach a situation of mutual satiation maintaining the observed gap. What happens in between the two observables can be assigned the name "black box" according to Kornai (1971). This is schematically represented in Fig. 2. The macroaggregative mismatch as perceived by microagents express mutual satiation trades.

Perception of Macroaggregative Mismatch → $\boxed{\text{Black Box}}$ → Expression of Mutual satiation Trades.

Fig. 2

The non-Walrasian structure is a structure applicable to the events occurring within the black box. The subjective construct hence is the construct hence is the construct conceptualising the formal representation of what is unobservable within the black box.

The subjective construct then starts with perception of the individual microagents of a macroaggregative observable. This perception by an individual agent is translated into effective demand of the individual agent. The microagents are all assumed to be identical. The buyers have identical preference patterns and the sellers face identical production possibility choices. The microagents then transform the effective demands into mutual satiation trades. The macroaggregative mutual satiation trade is then again the aggregative counterpart of the partial trades. So after expression of mutual satiation trades there is again a perception from microexpression to macroaggregates. So there is clearly an involvement of two-way preception which can be represented in terms of a schematic deagram in Fig. 3.

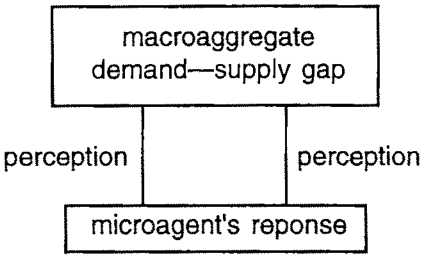


Fig. 3

The construct then of course has a means to relate these perceptions at two levels. The two perceptions must be felt in a similar way to have a proper microfoundation of a macroeconomic problem. In other words, the construct must allow symmetrical treatment of the two perceptions.

Leaving apart the initiation and rounding up of the entire process, we now take up the process in-between i.e. how the micro-agent's come to the expression of their trades of mutual satiation. The first feel of macroaggregate mismatch is being carried through perceived constraints. The perceived constraints are governed by notional demands at fixed prices. The buyer as well as the seller of a particular commodity is being governed by perception of a similar order. We have the same constraint perception rule for the two categories of microagents. The perception of the two agents participating in the exchange get expressed in their effective demands. There exists two paths corresponding to the time-series of perceptions of the two agents. One can be called the buyer's perception schedule and the other, the selller's perception schedule. The various points on the two schedules is a pair representing a contract. The two elements of the pair correspond to two points on the two paths respectively. Since effective demands exist at each point of time for all commodities for all microagents, such pairs exist. The two paths then speak of a process of mutual contract and recontract.

In the entire contract - recontruct process, the movement from one contract point to another recontract point owes itself to perceived constraint. A significant feature of perceived constraint is information received at the microlevel from macroaggregative data. The contract - recontract process can also incorporate certain extreme cases of perception like coexistence of macroaggregative excess supply with microlevel excess demand.

Thus placed, the entire contract - recontract process hinges on two distinct agents with respect to a single commodity. The positions of the two microagents are not symmetrical with respect to the contract - recontract process. In a situation of excess demand, we have a seller's market. The seller is placed definitely in a stronger position than the buyer. By the rule that the plans of the plans of the short side will be realised, the seller knows that he will be realise his plans. This is not so with the buyer. Information flow from macroaggregate to microlevel is entirely under the control of the seller in this situation. Uncertainty faced by the seller is much less than the uncertainty faced by the buyer. This implies that the buyer and the seller are placed entirely in an asymmetric position with respect to the contract - recontract process.

Asymmetric position of the two microagents to a process and two-way perception macroaggregative to microagents then form the subjective construct.

3. Applied Mathematical Method

The mathematical formalisation starts with the presumption of a gap between desired and actual transactions. Each individual "i" is represented by his representative utility function. The optimisation process yields the excess demand function Z_i^J (p) of individual "i" for commodity "j" given price vector of the commodities as "P". The individual who is on the long side of the market for a particular commodity faces a constraint. This is perceived constraint according to Benassy (1978), rationing schemes on purchases and sales according to Dreze (1975). By whatever name one may choose to call this constraint, perceived constraints or rationing schemes Z_i^J (p) is acted upon and transformed under this constraint. Hence the structure involves a sequence

$Z^J (p)^t = [Z_i^J (p); i = 1, 2, 3,....m]$

where Z^J (p) is the effective demand vector for commodity "j" at prices set at "P".

Questions on the problem are then answered by applying mathematical method on the axiomatic sturcture. The first question the structure faces is the question of existence. An existence question can be tackled only after defining an "equilibrium". Benassy (1978) defined a k-equilibrium. Younes defined a Younes equilibrium in terms of a terminating point in the sequence. The equilibrium represents a situation when all the individuals are mutually satiated.

The kind of mathematical analysis associated with a static non-Walrasian exchange problem is typical of the existential stage identified by mathematicians in the history of mathematical ideas. The existential stage stresses on the question of existence of solution rather than solutions themselves.

The particular theorem that has been used extensively in fixed point theorem. The basic theorem on fixed points which is being used is Brouwer fixed point theorem. The theorem asserts that if S is a compact, convex subset of Euclidian space, then every continuous function mapping K into itself has a fixed point. The theorem is applied to effective demand sequences. The fixed point has been identified to be a situation of mutual satiation. Fixed point is symbolising convergence of sequences $[Z^j\ (p)^t]$ where $[Z^j\ (p)^{t+1}]$ denotes the effective demand vector for commodity "j" at "t". The point of convergence has been identified with the "equilibrium" concept of non-Walrasian system.

To apply the fixed point theorem, one prerequisite is to have compact, convex domain and range. Either these must be assumed or constructed. Gale and Mas-Colell (1975) have shown that the restriction of the domain and range to compact, convex sets does not affect equilibrium. Hence we restrict our analysis to the intuitions carried by a fixed point rather than to the intuitions carried by convex, compact domains and ranges.

The intuition carried by the fixed point theorem is of an identity. The domain and range of the relevant mapping is to be identical. In the present application, domain and range of the effective demand sequences are to be identical. The sequence is schematically represented in Fig 4.

Fixed Point Mapping

Fig. 4

The identity of the domain and range does not necessarily mean static position. The existence of fixed point is guaranteed by mapping of a set into a set identical to itself. Except attainment of the fixed point, i.e. when a point gets transformed into itself under the mapping, there is no reason to assert that a particular state gets transformed to itself. The mapping of the point continues in a time sequence till attainment of the fixed point. Attainment of the fixed point then implies convergence of a process into a state. Each step in the process is characterised by mapping of a set into another point belonging to an identical set. The visualisation of a function is always that of a position or state into another state. Continued mapping represented by a sequence involves a process spread over time. The culmination of the process is in the form of a state. Hence fixed point implies culmination of a process of mapping a set into a set identical to itself.

Hence the process described intuitively by existence of a fixed point is a process when a state is being transformed to a state in an identical domain. Domain-range identity implies symmetrical position of the points with respect to each other. When y = f(x) where y belongs to S, x belongs to S and f represents a mapping; "x" and "y" are located symmetrically with respect to each other. We can say two points are located symmetrically with respect to each other if both are sustained under the same conditions and the pre and post conditions of the two points are similar. This is so as belonging to identical sets intuitively implies remaining within similar states.

The method of fixed point theorem implies existence of a process of mapping symmetrical positions into itself; which gets culminated in attainment of a fixed point.

4. Compatibility Exercise

Section 2 indicates the intuition implied by subjective construct of the problem. Section 3 derives the intuitions carried by the applied mathematical method. Now the basic compatibility exercise follows.

The contract - recontract process implied by the subjective construct involves another inherent process. This is the process of perception of a macroaggregate by microagents and perception of the microbehaviour at the macroaggregative level. The latter process is implied by the former. A contract is transformed into a recontract by transforming effective demands due to existence of perceived constraints. The effective demand sequence implies the process of contract and recontract. The sequence of perceived constraints implies a process of perception of macroaggregate by microagents and vice-versa.

The two processes are represented by two sequences – one of effective demands and other of perceived constraints.

$Z^J(P)$ and $\tilde{Z}^J(P)$

represent the effective demand and perceived constraints sequences faced in the market for jth commodity at price vector "p". The existing literature does not prove convergence of $Z^J(P)$ and $\tilde{Z}^J(P)$ separately. Convergence of $Z^J(P)$ is implied by convergence of $\tilde{Z}^J(P)$ as $Z^J(P)$ and $\tilde{Z}^J(P)$ are related by functional relationship. The two processes implied in the subjective construct are well represented by the two processes implied by applied mathematical method.

The mathematical analysis hinges on identical positions of domain and range. This identical position again implies that $Z^J(P)^t$ and $Z^J(P)^{t+1}$ are symmetrically situated with respect to each other. A particular position characterised by $Z^J(P)^t$ involves elements both from the buyer's path and seller's path. The buyer's path and the seller's path cannot be symmetrically situated with respect to each other. The extent and nature of asymmetry between the two paths need not be identical over time. Domain-range identity implies identity of asymmetry in the two consecutive time periods. This maintenance of the nature and extent of asymmetry at a single position, is imposing stronger restriction than the subjective construct can afford. In this place, the applied method distinctly deviates from a frame of subjective construct. The two are incompatible.

Compatibility between the two on the ground mentioned above can only be guaranteed at the position of attainment of fixed point. This is the position when trade of mutual satiation is attained. If any state is attained it is repeated. In this case the nature and extent of asymmetry can be maintained at an identical position. But the grounds for compatibility then becomes attainment of the fixed point or the solution. However even in the situation when compatibility question does not arise, existence of the solution is guaranteed only in presence of continuity of sequences. The presumption of continuity is limited when it is applied to observable entities. In case of observable entities, what can be observed is always observed. What cannot be observed does not exist. In this illustration, the macroeconomic aggregates are observables.

Distinction is to be made in this context between the intuition of contiunity in material sciences and social science e.g economics. In material sciences there is a limit to finenes of our powers of observation and the continuity property is a tool to tackle this limitation. If motion of a particle is observed at two different points the particle must have moved in between the two points even if its motion is not observable. Hence formalisation of continuity helps individuals in conceiving the abstractness of motion. So continuity is neither contrary nor counter-intuitive to the subjective constructive in material sciences. In social sciences, it is not true that a part of observables escape our observation. The sequence of mutual trades must be such that it consists of certain distinct attainable points. But the set of attainable points may be discrete and always not continuous. If the continuity is assumed, then the applied mathematical method imposes a structure which is non-existent, in the subjective construct. Hence there is no guarantee that the solution or the fixed point exists within the structure presented by the subjective construct. The fixed point theorem guarantees fixed point to exist in the structure imposed by the mathematical method, but does not guarantee existence of the solution within the rest observables in the economic structure. Hence the doubt of reaching a solution of mutual satiation is a very genuine doubt.

5. Towards An Alternative

The discussions in the preceding sections has led us to a controversial situation. I'm afraid we have entered into an area where opinions may broadly differ. Fixed point theorem used as a mathematical method in the subjective construct of non-Walrasian exchange problem has created certain problems. The problem is two-fold. The used method has some omitted features of the subjective construct. The method also expands the feasible area of operation of subjective construct. The deficiencies lead to the formation of an alternative mathematical structure derived from the subjective construct.

We start with the deficiencies of the old method. The first pointed out deficiency is inability to perceive asymmetry of the buyer and seller with respect to a market characterised by non-Walrasian exchange. The second one is incorporation of contract and recontract in the feasible set of operations, which may not be achieved in reality.

Let a buyer "i" of commodity "h" be represented by B_{ih}. B_{ih} is completely defined not only by his effective demand Z_{ih}, but also by the fact how he is placed in the market relative to the seller, denoted by S_{ih}. We indicate B_{ih} and S_{ih} by their position vectors $\langle P_{ih}|$ and $|P_{ih}\rangle$. Position has got two components–one indicating the distance from the market and the other indicating effective demand. The distance from the market is a concept similar to the tension of the buyer's aspiration according to Kornai (1971). Kornai (1971) defined the tension as follows:

$$\epsilon_{ih}^{Bt} = \alpha_{ih}^{t} - \bar{Z}_{ih}^{t}$$

ϵ_{ih}^{Bt} is the tension of buyer i for commodity h in time t

α_{ih}^{t} is the buyer's aspiration to buy commodity h in time t.

$\bar{Z}_{ih}^{t}$ is the actual transactions of commodity h intended by buyer i in time t.

Distance of the buyer denoted by d_{ih}^{Bt} can be defined by

$$d_{ih}^{Bt} = g\,(\epsilon_{ih}^{Bt}) \qquad \text{where } g' > 0$$

More is the tension, greater is the distance from the market with respect to the seller. The same analysis follows with respect to the seller. Exchange can be symbolised by the product.

All exchanges are not real. Only certain discrete exchanges form the feasible set of exchanges. The feasible set of exchanges can be constructed by applying a linear operator on the product as $c\langle P_{ih} | P_{ih} \rangle$. At the point of termination of the contract and recontract process, we reach a situation where both the buyer and the seller are in a symmetrical position with respect to the market. In a symmetrical vector we have the product defined as a matrix "g". The linear operator "c" corresponds to observable points of time. The matrix $c\langle P_{ih} | P_{ih} \rangle$ can be represented in two - dimensional plane as follows:

Let $\langle P_{ih} | = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$ where $a_1 b_1$ are the components of the vector $i = 1,2$

So $\langle P_{ih} | P_{ih} \rangle = \begin{bmatrix} a_1 a_2 & a_1 b_2 \\ a_2 b_1 & b_1 b_2 \end{bmatrix}$

In the structure a_i's denote position and b_i's denote effective demand. We $a_i b_i = 0$. This is so as the distance of one agent acted upon effective demand of the same agent is assumed to yield no effect. Exchange is embodied in the expressions $a_1 b_2$ and $a_2 b_1$.

The equilibrium is charatensed by a situation

$a_1 b_2 = 1$

$a_2 b_1 = 1$

Hence "g" $= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

represents an equilibrium matrix. Whether an equilibrium is reached or not is determined by

whether $\langle P_{ih} | P_{ih} \rangle$ = "g" or not.

Although the above analysis has been done with respect to $\langle P_{ih} | P_{ih} \rangle$ it can be easily extended to $c\langle P_{ih} | P_{ih} \rangle$. Again the entire analysis has been done with respect to a single buyer and a single seller. It can be extended to a situation of many buyers and many sellers. This is done by simple aggregation of entities over individuals. The subjective construct allows settling of partial trades as preconditions to settling of macrotrade balance. The structure therefore can be utilised to deal with existence, stability, and comparative static questions.

BIBLIOGRAPHY

Barro, R. J and Grossman, H. I. (1976) Money, Employment and Inflation, London and New York: Cambridge University Press

Benassy, J P (1978). A Neo-Keynesian Model of Price and Quality determination in G. SCHDIAUER ed. Equilibrium and Disequilibrium in Economic Theory, (1978), Boston: Reidel

Clower, R. W (1965) The Keynesian counter revolution. A Theoretical Appraisal in F H. HAHN and F. P. R BRECHLING (ed.), The theory of Interest Rates, London. Macmillan.

Dreze, J. H (1975). "Existence of an exchange equilibrium under price rigidities" In International Economic Review, Vol. 16

Gale, D and Mas-Colell (1975) "An Equilibrium Existence Theorem for a general model without ordered preferences" in Journal of Mathematical Economics, Vol. 2.

Grandmont, J. M. (1982): "Temporary general equilibrium theory" In K. J Arrow and M D Intriligator (ed.) Handbook of Mathematical Economics, Amsterdam North-Holland.

Grandmont, J. M and Laroque, G (1976): "On Keynsian Temporary Equilibria" In Review of Economic Studies, Vol 43

Hahn, F H. and Negishi, T (1962) "A theorem on non-tatonnement stability", in Econometrica, Vol 30

Kornai, J (1971): Anti-Equilibrium, Amsterdam: North-Holland.

Malinvaud, E (1977): The Theory of Unemployment Reconsidered. Oxford. Blackwell.

Malinvaud, E and Younes, Y (1977): "Some new concepts for the microeconomic foundations of macroeconomics" in G HARCOURT (ed) The Microeconomic foundations of Macroeconomics, (1977), London, Macmillan.

Uzawa, H (1962) "On the Stability of Edgeworth's barter process" in International Economic Review, Vol 3

Joint Provision Of Public Goods*

Asani Sarkar

This paper considers how two agents will jointly provide for a public project when the cost of the project is initially unknown to both. Any agreement between the agents regarding the funding of the project must take into account, first, the possibility that each agent will have an incentive to free-ride on the other's contributions and, second, the effect of cost uncertainty.

In a recent paper Admati and Perry [1] (AP from now on) examined the free-rider problem when the project cost is perfectly known to both agents. Agents move in alternate periods to make contributions towards meeting the cost of the project. Once made, these contributions are sunk. The project is completed when its full cost is borne and agents receive (identical) benefits from the project only when it is completed. To show how the free rider problem might arise in the context of this model, AP provides the following example. Suppose the agreement requires each agent to provide exactly half the project cost and the first agent keeps his side of the bargain. Since this contribution is sunk, the second agent can credibly ask to renegotiate the contract in his favour. Anticipating this, the first agent will refuse to contribute before agent two has paid his share.

AP show that if players are impatient and the disutility from making contributions is a strictly convex function, then there will be a (generically) unique number of periods in which the project will be completed in equilibrium if costs are not too high. To arrive at this equilibrium, AP first define a sequence of contributions $\{R_q\}_{q=1}^{\infty}$ where R_q is the maximum amount which makes player one indifferent between (1) contributing R_q today and having the project completed in q periods with player two contributing R_{q-1} in period 2, player one contributing R_{q-2} in period 3 and so on down the R_q sequence and (2) contributing 0 today and having the project completed in (q+1) periods by each player following the R_q sequence from period 2. It is important to note that R_q is independent of project costs. Further, player one's disutility from contributing R_q

---

*I am grateful to Richard Kihlstrom, George Mailath and Motty Perry for helpful comments I also thank seminar participants at the University of Pennsylvania.

in period 1 is equal to his disutility from having the project completed in (q+1) rather than q periods i.e., suffering a one-period delay relative to period q.

Suppose K satisfies $R_1 < K < (R_1 + R_2)$. Then in equilibrium, player one will contribute $(K-R_1)$ in period 1, player two will contribute R_2 in period 2 and the project will be completed in exactly two periods. To interpret this equilibrium, note that $(K-R_1) < R_2$ by hypothesis. Consider a deviation by player one to 0 in period 1. This delays project completion by one period. If W(x) is the disutility of contributing x, then from my earlier remark this delay costs player one disutility equal to $W(R_2)$ which is greater than $W(K-R_1)$, the utility gain from contributing 0 currently. If $K = (R_1 + R_2)$, it is clear that an additional equilibrium path will be $(0, R_2, R_1)$ since the utility gain and loss from contributing 0 rather than $(K-R_1) = R_2$ are exactly balanced. This result generalizes to any N such that K is nested between the sum of the R_q's over (N–1) and N, respectively. One rather useful observation is that, if players complete the project in q periods by following the AP sequence $(R_1, R_{q-1}, \ldots, R_1)$, any alternative sequence that delays project completion by more than one period will yield strictly lower payoffs to player one. This follows from my interpretation of R_q earlier.

The AP solution is not socially efficient since the equilibrium completion time may be greater than the socially optimal completion time. Additional inefficiencies may arise when costs are initially unknown. First, agents may withhold making contributions prior to the resolution of uncertainty, delaying project completion further. In fact, delay may occur in cases where completion is earlier than the first-best in the AP model. Second, agents may over-fund the project relative to the first-best. This will occur if the project is completed at high cost and later the "true" cost is realized to be lower.

I consider a variant of the AP model where the project cost is uncertain and may be high or low with some probability. Both agents are ignorant of the true cost at the onset but it is common knowledge that the cost will be revealed after a fixed number of periods L. It is assumed that agents always have the option of completing the project by contributing the high cost but if an amount equal to the low cost is contributed then the project will be delayed till at least period L. If the true cost is revealed to be low then the project will be completed in L periods. If it is high, then players must continue to make contributions until the full cost is met. One implication of the assumption is that if players have contributed a certain amount prior to period L and the realized cost turns out to be lower, the difference is lost.

To motivate this assumption suppose two partners of a firm must decide how to fund a project from which they both profit. The project can be completed with the help of a mainframe computer, which is currently available. However, a new powerful PC may arrive on the market in a year and it will be possible to complete the project with its help at lower cost. The partners' contract with the computer supply company allows them to buy the mainframe at any time once they pay for it. Alternatively, they may put down money for the PC which will be delivered in a year if the research bears fruit, in which case the project will be delayed for at least a year. When the partners bargain over their funding of the project, therefore, they face a trade off between finishing the project early at high cost and paying (possibly) less but at the cost of delaying project completion.

The project technology is considered to be low risk if the difference between the high and lost costs is small in the sense that, under perfect information about costs, the project would be completed in the same number of periods irrespective of whether costs were high or low. In fact, the only difference in the two perfect information equlibrium paths would be in the period 1 contribution, which would be lower for the low cost.

Now suppose H is the number of periods it would take for the project to be completed if costs were high for sure.(From AP it is known that H exists and is unique.) Consider the case where $H \geq L$. From my earlier remarks, it is clear that the project will be completed in max (H,L) = H periods if costs were low for sure, given the assumption on how the two costs affect outcome. So, under perfect information, both equilibrium paths would involve finishing the project in H periods with only the period 1 contribution differing. Since costs will be revealed before the project is completed, under uncertainty players will expect to "save" on their contributions by an amount exactly equal to the cost-difference if the low cost is revealed in period L. Further, this "savings" is only available in period L since once the contributions are adjusted for this difference, the high and low contributions schedules are identical by assumption. It turns out that this savings is maximized if the contributions schedule is high because the disutility function is convex and the equilibrium contributions sequence is increasing over time. In periods other than L the logic of the AP result implies that players can do no better than contribute the high amount since any smaller period 1 contribution (say) will delay the project by at least one period. So the equilibrium path is: players contribute the high amount until costs are revealed and then follow the AP sequence.

There is another equilibrium path where, if the probability of the low cost being realized is high enough, players choose to save in period (L–1) rather than L by contributing the low amount. This is a surprising result since the lower cost technology is not actually available in period (L–1). Therefore, if players' anticipation is wrong and realized costs turn out to be high, they must reswitch to funding the higher cost technology again in period L.

If H < L, then players have the option of completing the project before costs are revealed by following the high contributions schedule. However, this means giving up the savings that could have been obtained if the project had extended for L periods. Indeed, two types of strategies are seen to be possible in equilibrium. If the probability of the high cost being realized is sufficiently high, the (HS) strategy is followed where HS involves early completion of the project in H periods by following the high contributions schedule. For a low probability of the realization of the high cost, the alternative (DC) strategy involves contributing nothing for (L–H) periods and then following HS. In this case the project is completed in L periods. Thus, players delay the project by just enough to access the expected savings in period L. If the project is extremely valuable (formally H ≤ L/2), players are not willing to delay and HS is unique irrespective of probabilities.

An interesting variant of the DC strategy, called DCL, is obtained when players choose to save in period (L–1) by contributing the lower amount. This introduces an additional delay of one period relative to HS. Therefore, as the probability of the high cost being realized decreases, a situation may arise where players switch from HS to DC to HS again and finally to DCL. This happens if the additional one period delay increases players' disutility sufficiently to make DCL more costly to follow relative to HS.

Analogous to the low-risk case, the project technology is deemed to be risky if the difference between the costs is large. In patrticular, the low cost is so low that the project can be completed in one period (although players must still wait till period L to realize its benefits). With risky technology, one equlibrium path involves players not contributing till costs are revealed. The reason is that, since immediate completion is assumed if the low cost is realized in period L, any contributions made prior to L is made "unnecessary." As before, HS is unique for H ≤ L/2.

Finally, under an additional assumption, it is shown that the project may not be completed in equilibrium. If the low cost is too high, the project is obviously not completed. But suppose that the low cost is low while the high cost is so high that the project would not be completed under perfect information. Suppose L = 2. Then the equilibrium involves contributing either zero or the low amount in period 1. Suppose some small positive contribution is made compulsory in period 1 (interpret this as the minimum investment needed to learn the true costs in period 2). Now the period 1 contribution is always positive but players realize the project's benefits only if the cost is revealed to be low in period 2 (by hypothesis). When the expected benefits are below the disutility of the period 1 contribution, players' expected utility from participation in the project in the project is negative. Hence the result.

The paper is organized as follows. Section 1 reviews the AP model while section 2 presents the bargaining model with uncertain costs. Sections 3 and 4 derive the results for low-risk technology. In section 5, I describe the equilibrium when the technology is risky, section 6 characterizes the properties of the equilibria and section 7 concludes.

Section 1. The AP Model

There are two identical players, one and two, who move in alternate periods to finish a project that costs K. V is the value of immediate completion to each player. If c_i^t is the contribution of player i at time t and TC_i^t is the total contribution of player i from the beginning of the game up to period t then the game ends in T periods if the following condition holds:

$$TC_1^T + TC_2^T \geq K \tag{1.1}$$

A history of length T is $(\{c_1^t\}_{t=1}^T, \{c_2^t\}_{t=1}^T)$. $c_i^t = 0$ if it is not the turn of player i to move in period t. A strategy for player i is a function from each history after which it is his turn to move to a contribution level. Players are assumed to be impatient i.e., they have a time discount factor $\delta < 1$. An outcome is $(T, \{c_1^t\}_{t=1}^T, \{c_2^t\}_{t=1}^T)$ where $T = \infty$ if the project is not completed. A payoff for each player is defined by

$$U_i(T, \{c_1^t\}_{t=1}^T, \{c_2^t\}_{t=1}^T) = \delta^{T-1} V - \sum_{t=1}^{T} \delta^{T-1} W(c_i^t) \tag{1.2}$$

W (.) is the effort function. It is assumed to be increasing and strictly convex with W (0) = 0.

D1.1

The extensive form game described above will be known as the AP game with respect to K.

Admati and Perry set out to find the sub-game perfect equilibrium (SPE from now on) of this game. They define a sequence of critical contribution levels $\{R_q\}_{q=0}^{\infty}$ where

$$R_0 = 0 \tag{1.3}$$

$$R_1 = W^{-1}\{V(1-\delta)\} \tag{1.4}$$

$$R_n = W^{-1}\{V.\delta^{2n-3}(1-\delta^2)\} \text{ for } n \geq 2 \tag{1.5}$$

R_1 is the maximum amount such that player i is indifferent between (i) contributing R_1 and completing the game today and (ii) investing 0, letting player $j \neq i$ contribute R_1 and complete the project tomorrow. By recursion R_n is a level of contribution that makes player i indifferent between the following options:

(i) contribute R_n, so that $\sum_{q=0}^{n-1} R_q$ is left to be invested and from then on player $j \neq i$ contributes R_{n-1}, i contributes R_{n-2}, etc. until the project is completed.

(ii) contribute 0, so that $\sum_{q=0}^{n} R_q$ remains to be invested and from then on player $j \neq i$ contributes R_n, i contributes R_{n-1}, etc. until the project is completed.

Consider the contributions sequences implied by (i) and (ii). Player one's disutility from contributing 0 instead of R_q in period 1 are : (1) a one-period delay in project completion time and (2) a change in the sequence of contributions. By definition, the net disutility which I will call the delay cost must be equal to $W(R_q)$. Also, the R_q's themselves are just numbers and so idependent of K. Remark 1.1 summarizes these observations.

Remark 1.1

(a) $W(R_q)$ measures the disutility of a one-period delay in project completion relative to period q. (b) The contribution levels R_q are independent of project costs. (c) For $q \geq 2$, $W(R_{q+1}) = \delta W(R_q)$ and so is decreasing in q. $R_2 \gtreqless R_1$ according as $\delta \gtreqless \delta^*$ where $\delta^* = 0.62$.[1]

Let N^* be the unique integer satisfying the following condition:

$$\sum_{q=0}^{N^*-1} R_q < K \leq \sum_{q=0}^{N^*} R_q \tag{1.6}$$

D 1.2

The SPE of the AP game with respect to K is given as follows.

(i) Suppose $\sum_{q=0}^{\infty} R_q > K$

(a) If $\sum_{q=0}^{N^*} R_q > K$ then the unique equilibrium path is

$c_1^1 = K - \sum_{q=0}^{N^*-1} R_q, c_i^t = R_{N^*-t+1}$ for $1 < t \leq N^*$ if it is player i's turn to move in period t. The project is completed in N* periods.

(b) If $\sum_{q=0}^{N^*} R_q = K$ then there are two equilibrium paths. In addition to (a) above there is another path with $c_1^1 = 0$ and after that the game continues as in (a) with $c_2^2 = K - \sum_{q=0}^{N^*-1} R_q$

(ii) $\sum_{q=0}^{\infty} R_q \leq K$, in which case $c_i^t = 0$ for all i and all t. The project is left incomplete.

Remark 1.2

Suppose $N^* \geq 2$ player one's equilibrium payoff in part (ia) is $\delta^{2N^*-3} V - W - W\left(K - \sum_{q=0}^{N^*-1} R_q\right)$ and it is $\delta^{2N^*-1} V$ in part (ib) of D1.2. Player two's equilibrium payoff in (ia) is $\delta^{2N^*-2} V$.

Section 2. Bargaining with Uncertain Costs

The AP model described in section 1 is modified by assuming that the project cost K is unknown to both players. However, it is common knowledge that K has the following distribution:

$$\begin{aligned} \tilde{K} = K_h & \text{ with probability } (1-p) \\ K_\ell & \text{ with probability } p \end{aligned} \tag{2.1}$$

It is also common knowledge that K will become known after L periods. Agents enjoy the benefit of the project as soon as it is completed. The project is deemed to be completed in period t when

$$\begin{aligned} &\text{either } TC_1^t + TC_2^t \geq K_h \text{ and } t < L \\ &\text{or } TC_1^t + TC_2^t \geq K_j \text{ with } K = K_j \\ &\text{and } t \geq L \text{ for } j = h, \ell \end{aligned} \tag{2.2}$$

Eq. (2.2) states that the players have current access only to the expensive technology. Thus agents always have the option of contributing the high cost and completing the project, even before period L. However, if they choose to contribute the lower amount they have to wait till period L to find out whether in fact the low cost has been realized. If it turns out that the cost is K_h then the game must continue until the full cost is borne by the players. Players, therefore, face a tradeoff when deciding upon their investment strategies. If the project is sufficiently valuable to them they can invest the high amount and complete the project earlier. Or, they can choose to delay and gamble that they may have to contribute a lower amount in period L.

Let c_{ij}^t be player i's contribution in period t if it is known that the true cost is K_j. Let c_i^t be this contribution when the true cost is not known. Three types of histories can then be discerned. Let S (n) be a history of length n. Suppose the project is completed before L periods. Then :

$$S(T) = \left(\left\{ c_1^t \right\}_{t=1}^{T} \left\{ c_2^t \right\}_{t=1}^{T} \right) \text{ Where } T < L.$$

Suppose the project is still incomplete at the end of period L-1. Define $T_j \geq L$ as the smallest number of periods in which the total contributions of the two players exceed K_j, for $j = h, \ell$. Note that $T_h \geq T_\ell$. Then :

$$S(T_j) = \left(\left\{c_1^t\right\}_{t=1}^{L-1}, \left\{c_2^t\right\}_{t=1}^{L-1}, \left\{c_{1j}^t\right\}_{t=L}^{T_j}, \left\{c_{2j}^t\right\}_{t=L}^{T_j}, K_j \right), j = h, \ell.$$

If the project is completed in T < L periods, then an outcome is $\left(\left\{c_1^t\right\}_{t=1}^{T}, \left\{c_2^t\right\}_{t=1}^{T}, T \right)$ and player i receives :

$$U_i\left(T, \left\{c_1^t\right\}_{t=1}^{T}, \left\{c_2^t\right\}_{t=1}^{T} \right) = \delta^{T-1}V - \sum_{t=1}^{T} \delta^{T-1} W\left(c_i^t\right) \qquad (2.3)$$

If the project is completed in $T_j \geq L$ periods then an outcome is $O_j = \left(\left\{c_1^t\right\}_{t=1}^{L-1}, \left\{c_2^t\right\}_{t=1}^{L-1}, \left\{c_{1j}^t\right\}_{t=L}^{T_j}, \left\{c_{2j}^t\right\}_{t=L}^{T_j}, T_j, K_j \right)$ for $j = h, \ell$ and player i's expected payoff is :

$$U_i(O_h, O_\ell) = (1-p)\left[\delta^{T_h-1}V - \sum_{t=L}^{T_h} \delta^{t-1} W(c_{ih}^t)\right]$$

$$+p\left[\delta^{T_\ell-1}V - \sum_{t=L}^{T_\ell} \delta^{t-1} W(c_{i\ell}^t)\right] - \sum_{t=1}^{L-1} \delta^{t-1} W(c_i^t) \qquad (2.4)$$

Section 3. The Solution for Low-Risk Technology: H>L

Let H be the unique integer which satisfies:

$$\sum_{q=0}^{H-1} R_q < K_h \leq \sum_{q=0}^{H} R_q \qquad (3.1)$$

In this section I will concentrate on the case where $H \geq L$.

Section 4 deals with H < L. Further, I will assume that the technology is not too risky in thefollowing sense.

D 3.1

Define $k = (K_h - K_\ell)$. The technology is low risk if:

$$\sum_{q=0} R_q \leq K_\ell < K_h \leq \sum_{q=0} R_q \tag{3.2}$$

which implies that $k \leq R_H$. Since Rq is decreasing in q for $q \geq 2$, (3.2) implies that project with longer equilibrium completion times will have smaller cost differences.

Define the following contribution and probability functions:

$$C_j(H) = K_j - \sum_{q=0}^{H-1} R_q \quad \text{for } j = H, L \tag{3.3}$$

$$C_k(H) = k - \sum_{q=0}^{H-1} R_q$$

$$q^1(H) = 1 - \frac{[W(C_h) - W(C_\ell)]}{W(R_h)} \tag{3.4}$$

C_j and C_k are functions of H, but this dependence will be suppressed for notational simplicity. Also define the following contributions sequences:

(HS) $C_1^1 = C_h$; $\{R_{H-1}, R_{H-2}, \ldots, R_{H-L+2}\}$ up to period (L-1); in period L

$\{R_{H-L+1}$ if $\tilde{K} = K_h$ or $(R_{H-L+1} - k)$ if $\tilde{K} = K_\ell\}$;

$\{R_{H-1}, \ldots, R_1\}$ from period (L+1) till H.

(LS) $C_1^1 = C_\ell$ and then follows (HS) from period 2. That is, $\{k, R_{H-1}, \ldots, R_{H-L+3}\}$ up to period (L–1); R_{H-L+2} if $\tilde{K} = K_h$ or $(R_{H-L+2} - K)$ if $\tilde{K} = K_\ell\}$ in period L; $\{R_{H-L+1}, R_1\}$ from period (L+1) till period (H+1).

(OS)$C_1^1 = 0$ and then follow (HS) from period 2.

Observe that when $L = 2$ and $p = q1(H)$, player one is indifferent between the high contribution schedule (HS) and the low contributions schedule (LS). To see this, rewrite (3.4) as $[1-q1(H)]\ W(R_H) = W(c_h) - W(c_l)$. The right-hand side (heneforth RHS) of the expression is the gain in utility from funding the cheaper technology C_l in period 1. The left-hand side (LHS from now on) is the expected disutility due to the one-period delay in project completion if C_l is paid in period 1 but the high cost is realized in period L. $[1-q1(H)]$ is the probability of K_h being realized while $W(R_H)$ is the disutility from having the project completed in H rather than H+1 periods.

The equilibrium is derived with the help of a lemma. Let G(T,N) denote the sub-game when the project can be completed in N periods using the expensive technology and the true cost will be realized after T periods.

Remark 3.1

In G(O,N) (i.e., all sub-games starting from period L), the AP result defined in D1.2 can be applied directly since from period L costs are known with certainty.

The following lemma says that, after any history such that the players are in a G(2,H) sub-game and $H \geq 2$, players may switch from (HS) to (LS) if p is high enough.

Lemma 3.1

Let $L = 2$ and $H \geq L$.

(i) If $P < q1(H)$, then the SPE path is (a) (HS) when $C_h < R_h$. The project is completed in H periods. (b) (HS) and (OS) when $C_h = R_h$. Completion is in (H+1) periods for (OS).

(ii) If $P > q1(H)$, the unique SPE path is (LS) and the project is completed in (H+1) periods.

(iii) If $P = q1(H)$, then the SPE paths are (a) (HS) and (LS) if $C_h < R_H$ and (b) (HS), (OS) and (LS) when $C_h = R_H$.

Proof

First, I show that C_h, C_ℓ and 0 are the only period 1 contributions that occur in any SPE. Consider $C_1^1 = x$. Then the amount of project cost remaining to be invested in period 2 is $(K_h - x)$ if $\tilde{K} = K_h$ or $(K_\ell - x)$ if $\tilde{K} = K_\ell$.

But $\sum_{q=0}^{H-1} R_1 < (K_h - x) \leq (\sum_{q=0}^{H} R_q - x) < \sum_{q=0}^{H} R_q$ and

$\sum_{q=0}^{H-2} R_q < (K_\ell - x) < \sum_{q=0}^{H-1} R_q$. So, applying the AP result, player one's contributions sequence from period 3 will be identical to (LS). Since $C_\ell < x$, x will never be contributed in period 1 in an SPE.

Similarly, consider $0 < y < c_\ell$. The amount of project cost remaining in period 2 is $(K_h - y)$ or $(K_\ell - y)$. Since $\sum_{q=0}^{H-1} R_q < (K_\ell - y)$ $< (K_h - y) < \sum_{q=0}^{H} R_q$, player one's contributions sequence from period 3 will be identical to (OS) and so $C_1^1 = 0$ will always be preferred to y in any SPE.

Next, compare (HS) and (OS). Player one gains $W(C_h)$ in period 1 from following (OS) but at the expense of delaying project completion by one period entailing a disutility of $W(R_H)$ according to Remark 1.1. When $C_h < R_H$, (HS) is preferred; when $C_h = R_H$, player one is indifferent between (HS) and (OS).

Finally, compare (HS) and (LS). As mentioned previously, when $p = q1(H)$ player one is indifferent between (HS) and (LS). When $p < p1(H)$, $[1-p]\ W(R_H) > [W(C_h) - W(C_\ell)]$ and so (HS) is preferred. The rest of the lemma follows immediately.

Lemma 3.1 suggests that a path in which players contribute according to (LS) from period (L–1) should be a candidate SPE profile. So define these contributions sequences:

(HLS) Follow (HS) up to period (L–2) and (LS) from (L–1).

(LLS) Follow (LS) up to period (L–2) and (LS) from (L–1).

(OLS) Follow (OS) up to period (L–2) and (LS) subsequently.

Now suppose agents are making a choice between following (HS) and (HLS). Up to period (L–2) the paths are identical. So consider the sub-game beginning at period (L–1). The player who moves in that period gains $[W(R_{H-L+2}) - W(R_{H-L+2}-k)]$ in period (L–1) by switching from (HS) to (HLS). If costs are revealed to be low in period L, then the lower funding is justified and entails no delay. But if costs turn out to be high, then with probability (1–p) there is a disutility of $W(R_{H-L+2})$ suffered due to a one-period delay in project completion relative to period H. If these costs and benefits balance exactly when p = q2(H,L) then:

$$q2\ (H,L) = \frac{W(R_{H-L+2}-k)}{W(R_{H-L+2})} \tag{3.5}$$

The next proposition characterizes the equilibrium for $H \geq L$ and $L \geq 3$.

Proposition 3.1

Let $L \geq 3$ and $H \geq L$. Suppose the technology is low risk in the sense of D 3.1. Assume that k satisfies the lower bound defined below in (3.8) and $\delta < 0.62$ (which implies $R_1 > R_2$). Then:

(i) if L is odd and (a) $p \leq q^2(H,L)$ then the high contributions schedule (HS) is the unique SPE path. The project is completed in H periods. (b) If $p > q^2(H,L)$ and $C_h < R_H$, (HLS) is the unique SPE path. The project is completed in (H+1) periods if $\tilde{K} = K_h$ and H periods if $\tilde{K} = K_\ell$. When $C_h = R_H$, an additional SPE path is (OLS).

(ii) If L is even and (a) $p \leq q^2(H,L)$, then (HS) is the unique SPE path when $C_h < R_H$. In addition, (OS) is also a SPE path from $C_h = R_h$. (b) If $p > q^2(H,L)$, then (HLS) is the unique SPE path.

Proof

Consider L=3, $H \geq L$. If agents are choosing between (HS) and (HLS), it follows from the discussion prior to equation (3.5) that the choice will be :

(HS) $c_1^1 = C_h$ and then $\{R_{H-1}, (R_{H-2}$ if $\tilde{K} = K_h$ or $R_{H-2} - k$ if $\tilde{K} = K_\ell$), $R_{H-3}, \ldots, R_1\}$ from period 2 if $p \leq q^2(H,3)$; and

(HLS) $c_1^1 = C_h$ and then $\{R_{H-1} - k, (k$ or $R_{H-2})$, $(R_{H-2}$ or $R_{H-1}), \ldots, (R_1$ or $0)\}$ from period 2 if $p \geq q^2(H,3)$.

Suppose agents are choosing between (LS) and (LLS). $c_1^1 = C_\ell$ in both paths. In period 2, player two gains W(k) by deviating to (LLS) but loses (1−p) W (R_H) since the project is delayed by one period (relative to period H) if costs are high. If these gains and losses are offset at $p = q^3(H)$ then:

$$q^3(H) = 1 - \frac{W(k)}{W(R_H)} \tag{3.6}$$

Therefore agents select:

(LS) $c_1^1 = C_\ell$ and then $\{k, (R_{H-1}$ or $R_{H-1} - k), R_{H-2}, \ldots, R_1\}$ from period 2 till (H+1) if $p \leq q^3(H)$; and

(LLS) $c_1^1 = C_\ell$ and then $\{0, (k$ or $R_{H-1}), (R_{H-1}$ or $R_{H-2}), \ldots, (R_2$ or $R_1), (R_1$ or $0)\}$ from period 2 till (H+2) if $p \geq q^3(H)$.

Similar reasoning shows if $c_1^1 = 0$ agents select between (OS) and (OLS) in the following manner:

(OS) $c_1^1 = 0$ and then $\{C_h, (R_{H-1}$ or $R_{H-1} - k), R_{H-2}, \ldots, R_1\}$ from period 2 till (H+1) if $p \leq q^1(H)$; and

(OHS) $c_1^1 = 0$ and then $\{C_\ell, (k$ or $R_{H-1}), (R_{H-1}$ or $R_{H-2}), \ldots, (R_2$ or $R_1), (R_1$ or $0)\}$ from period 2 till (H+2) if $p \geq q^1(H)$.

The following observations indicate that (HS), (HLS), (LS), (LLS), (OS) and (OLS) are the candidate SPE strategy profiles: (1) From period 2, the sub-game is G(2,N) with $N \geq 2$ so that Lemma 3.1 can be applied directly. (2) The project will be completed since, by Lemma 3.1, once period 2 is reached, all possible equilibria involve completing the project. (3) The only relevant period 1 actions for player one are c_h, c_ℓ and 0 (the proof is similar to the one used in Lemma 3.1 and so is omitted). Further the probabilities are related in this way:

$$q^2(H,3) \geq q^3(H) > q^1(H) \tag{3.7}$$

To show (3.7), note that $W(k) < [W(c_h) - W(C_\ell)]$ since $W(x+y) > W(x) + W(y)$ and so $q^3 > q^1$. For $q^2(H,3) \geq q^3(H)$ it is necessary that:

$$\delta^2 W(R_{H-1}-k) - W(R_H) + W(k) \geq 0 \tag{3.8}$$

(3.8) is an assumption since it may be negative for small values of k when W (.) is strongly convex. Effectively, (3.1) places a lower bound on k. As an example, if $W(x) = x^2$, $k \geq \frac{2R_H\delta^2}{1+\delta^2}$ is needed. On the other hand, if either W is linear of $k = R_H$ then (3.8) is always positive.

Suppose $p \leq q^1(H)$. By observation, player one strictly prefers (OS) to (LS). Comparing (OS) and (HS), player one would benefit by $W(C_h)$ in period 1 by deviating to (OS) but would suffer disutility equal to $W(R_H)$ due to the one-period delay in project completion. In addition, under (HS), these is a "cost savings" equal to

$p\delta^2 [W(R_{H-2}) - W(R_{H-2}-k)]$ realized if $\tilde{K} = K_\ell$ instead of K_h in period 3. The amount of these "cost savings" is $p\delta^2 [W(R_{H-1}) - W(R_{H-1}-k)]$ under (OS). Since W(.) is strictly convex and R_q is decreasing in q, these "savings" are greater under (HS). So, for $C_h \leq R_H$, player one receives strictly greater utility under (HS).

Suppose $p \; \varepsilon \; (q^1(H), q^3(H)]$. Clearly (HS) is prefered to both (LS) and (OS). Further player one prefers (OS) to (OLS) since under (OS) player one derives utility equal to $\delta^{2H-1}V$ (see Remark 1.2) plus "cost savings" equal to $p\delta^2$ $[W(R_{H-1}) - W(R_{H-1}-k)]$ in period 3, but gets only $\delta^{2H-1}V - (1-p)\ \delta^2 W(k)$ under (OLS). To see this, replace k with 0 in period 3 of the (OLS) sequence. From Remark 3, player one's utility is $p\delta^{2H-1}V$ when $\tilde{K} = K_\ell$ and $(1-p)\delta^2\ (\delta^{2H-3}V)$ when $\tilde{K} = K_h$. When $p\ \varepsilon (q^3(H)\ q^2(H,3)]$, (HS) is again the preferred choice because (OLS) is preferred to (LLS) by observation. This proves (ia) of the proposition for L = 3.

Let $p > q^2(H,3)$ and compare (HLS) and (OLS). Player one's gain from a deviation to (OLS) is $W(C_h)$. If $\tilde{K} = K_\ell$, then the disutility from delay is $p.W(R_H)$. If $\tilde{K} = K_h$, without loss of generality put $k = 0$ in both sequences and apply Remark 1.2 to the sequence starting from period 3. This calculation yields an additional disutility of $(1-p)\ \delta^2 W(R_{H-1}) = [1-pw\,(R_H)]$. Part (ib) of the proposition follows immediately.

Next, the induction argument is used to generalize the proof. Suppose $p \leq q^2(H,L)$ and the result holds for L=T, $H \geq L$ and T is odd. Then (HS) is the unique SPE path. Consider L = (T+1) and $N \geq L$ where N = H+1. Note that L is even, so player one does not move in period L and does not realize any "cost savings." The candidate profiles are (HS), (LS) and (OS). (OS) is always preferred to (LS). Without the "cost savings," player one prefers (HS) to (OS) if $C_h < R_H$ and is indifferent between the two when $C_h = R_H$. If T is even $C_h < R_H$, then (HS) is again the unique SPE path but L is odd so that player one does obtain the "cost savings" which makes (HS) unique in the G (L,N) game also. When T is even and $C_h = R_H$, (OS) is also a SPE path in G(T,H). In the G(L,N) game, this induces an additional candidate profile with $C_1^1 = C_h$ and $C_2^2 = 0$. Clearly, since player one can always deviate to 0, this will never be a SPE path.

If $p > q^2(H,L)$, then (HLS) is the unique SPE in the G(T,H) game when T is odd and $C_h < R_H$. As always, (LLS) is ruled out in G(L,N). Comparing (HLS) and (OLS), observe that L is even and so player one moves in period (L−1) when the switch to the lower funding strategy is accomplished. This ensures "savings" equal to $\delta^{L-2}[W(R_{H-L+2}) - W(R_{H-L+2}-k)]$ under (HLS)and $\delta^{L-2}[W(R_{H-L+3}) - W(R_{H-L+3}-k)]$ in period L–1. Again, "savings" are greater under (HLS) because $R_{H-L+2} > R_{H-L+3}$ when $H \geq L$ and W(.) is convex, and so (HLS) is the unique SPE path. If T is even, then (HLS) is the unique SPE path in G(T,H). But L is odd, so player one does not move in period (L–1) and cannot enjoy any "savings" in that period. Therefore, for $C_h = R_H$, (HLS) and (OLS) yield the same utility to player one. Otherwise, for $C_h < R_H$, (LHS) is strictly preferred. This completes the proof.

Section 4. Equilibrium for H < L

When there is uncertainty about project costs, an important element determining equilibrium is the projected cost savings in period L if the low cost is realized. If the project is sufficiently valuable, however, players may decide to forego these cost savings and *always* complete the project prior to period L by funding the costly technology. Ex-post, therefore, it may turn out that players will have contributed more than the amount necessary to complete the project. In this section, I derive conditions under which there will be early completion *irrespective* of how probable the realization of the costly technology is.

When H < L, the high contributions profile (HS) take the following special form:

(HS) $\{C_h, R_{H-1}, \ldots, R_1\}$

Suppose H = 3 and L = 7. If players follow (HS) and finish the project in three periods, player one receives at least $\delta^5 V$ (see Remark 1.2). If player one contributes 0 in period 1 and delays completion by one period, he can do no better than receive $\delta^5 V$, as we know from the AP results. If the strategy is such that there is a two period delay, then player one will do strictly worse. To see this, note that with a two-period delay player one must finish the project in period 5 by contributing R_1 which yields $[\delta^4 V - \delta^4 W(R_1)] = \delta^5 V$. Since he must also contribute something in the earlier periods, he will get an amount less than $\delta^5 V$. However, if the project is sufficiently delayed so that the game reaches period L, player one can expect some cost savings if $K = K_l$ in period L. Depending on parameter values, these cost-savings may more than offset the cost of delay. This suggests the following alternative strategy: contribute nothing for (L–H) periods and then follow (HS) so that the project is completed in exactly L periods. Call this contributions sequence (DC) for delayed completion:

(DC) {0 for (L–H) periods; C_h in period (L–H+1); then R_{H-1}, ..., R_2 up to period (L–1); R_1 if $\tilde{K}=K_h$ or (R_1-k) if $\tilde{K}=K_l$ in period L}.

Note that the period L "savings" are only available to player one if L is odd. However, analogous to the (OSL) sequence in section 3, one can define a variant of (DC) in which player one switches to the low funding strategy (LS) in period (L–1) as follows:

(DCL) {0 for (L–H) periods; C_h , R_{H-1}, ...,R_3 up to period (L–2); (R_2-k) in period (L–1); $(C^2_L = k \text{ and } c^1_{L+1} = R_1)$ if $\tilde{K}=K_h$ or $C^2_L = R_1$ if $\tilde{K} = \overline{K}_\ell$}.

Under (DCL), player one can benefit from the "savings" in period (L–1) of $[W(R_2)-W(R_2-k)]$. Now suppose L is odd and player one must choose between the paths (HS) and (DC). First, suppose (L–H) is odd so that under (DC) player one contributes $\{R_{H-1}, R_{H-3}, ..\}$ from period (L–H+2). If player one deviates to (DC), the period 1 gain is $W(C_h)$ and the expected period L "savings" is $p.\delta^{L-1}[W(R_1)-W(R_1-k)]$. Against this, there is a delay of (L–H) periods relative to period H. Ignoring the period 1 contribution (which is already accounted for), player one derives utility worth $\delta^{2H-3}V$ under (HS) and $\delta^{L-H+1}.\ \delta^{2H-3}V$ under (DC).[2] So the cost of delay is $\delta^{2H-3}V(1-\delta^{L-H+1})$. Now, let (L–H) be even so that under (DC) player one contributes $\{C_h, R_{H-1}, ...\}$ from period (L–H+1). Therefore his disutility from the (L–H) period delay relative to (HS) is now $\delta^{2H-3}V - \delta^{L-H}\{\delta^{2H-3}V - W(C_h)\}$. So, if q^4 (H,L) is the probability value such that player one is indifferent between (HS) and (DC) in these two cases then :

$$q4(H,L)=\frac{\left\{\delta^{2H-3}V-W(c_h)\right\}\left(1-\delta^{L-H}\right)}{\delta^{L-1}\left\{W(R_1)-W(R_1-k)\right\}} \quad \text{if } (L-H) \text{ is even and L is odd} \qquad (4.1)$$

$$=\frac{\delta^{2H-3}V\left(1-\delta^{L-H+1}\right)-W(c_h)}{\delta^{L-1}\left\{W(R_1)-W(R_1-k)\right\}} \quad \text{if } (L-H) \text{ is odd and L is odd}$$

Similarly, when L is even, one can evaluate player one's choice between the paths (HS) and (DCL). The only difference with the previous calculation is that player one now moves in period (L–1) and "saves" $p.\delta^{L-2}\{W(R_2)-W(R_2-k)\}$. Define q5 (H, L) as the probability value such that player one is indifferent between (HS) and (DCL).

Then :

$$q5(H,L)=\frac{\left\{\delta^{2H-3}V-W(c_h)\right\}\left(1-\delta^{L-H}\right)}{\delta^{L-2}\left\{W(R_2)-W(R_2-k)\right\}} \quad \text{if } (L-H) \text{ is even and L is even} \qquad (4.2)$$

$$=\frac{\left\{\delta^{2H-3}V\left(1-\delta^{L-H+1}\right)-W(c_h)\right\}}{\delta^{L-2}\left\{W(R_2)-W(R_2-k)\right\}} \quad \text{if } (L-H) \text{ is odd and L is even}$$

Suppose H = (L–1) and L = 3. In this case, the path (DC) is identical to the path (OS) which was defined as the sequence $\{C_1^1=0$; then follow (HS) from period 2}. Similarly, (DCL) is identical to (OLS). If $C_h(H) = R_H$, then the AP result states that (HS) and (DC) yield the same payoff to player one in a full-information world. With uncertainty, however, player one also "saves" in period L under (DC) and so (DC) yields strictly higher payoffs. The implication is that, if $C_1^1=C_h(H)$ then in the period 2 sub-game G(L–1, H–1), $C_h(H-1) = R_{H-1}$ and so the optimal period 2 contribution is 0. In other words, when H = (L–1) the high contributions sequence (HS) can never occur in any SPE. The following lemma states that when H = (L–1) the SPE paths are either (DC) or (DCL).

Further (DCL) is chosen when $p > q^*$, where $q^* = q^2(H,H) = W(R_2-k)/W(R_2)$. The reason is: if $C_1^1 = 0$ the period 2 sub-game is G(H–1, L–1). In this sub-game, therefore, H = L and agents either follow the high contributions sequence (HS) or follow (HLS) by switching to the low funding strategy in period (L–1) if $q > q^2(H,L) = q^2(H,H) = q^*$ (see the discussion on q2 in section 3).

Lemma 4.1

Suppose H = (L–1). If (a) H = 3 and $C_h = R_H$ or (b) H > 3, then the unique SPE path is (DC) for $p \leq q^*$ and (DCL) for $p > q^*$ where $q^* = W(R_2-k) / W(R_2)$.

Proof See appendix.

What happens when H < (L–1)? First, notice that (HS) is the unique SPE path when H = 2 because player one gets at least $\delta^3 V$ here but since $L \geq 4$, player one's utility from (DC) or (DCL)–which involves project completion in period L or after—must be strictly less than $\delta^3 V$. In fact, this argument generalizes to any $H \leq L/2$, as will be shown later. Second, suppose H = 3 and L = 5. If $C_1^1 = C_h$, the period 2 sub-game is G(4,2) and so from the previous comment, (HS) is the unique equilibrium. If $C_1^1 = 0$, the period 2 sub-game is G(4,3) and Lemma 4.1 applies.[3] So for $p \leq q^*$, the comparison is between (HS) and (DC). If $q^4(3,5) < q^*$, then the discussion on q4 suggests that (HS) is preferred when $p \leq q^4(3,5)$ while (DC) selected when p belongs to the closed interval $[q^4(3,5), q^*]$. Since L is odd, q5 is not defined and (HS) is the unique SPE path for $p > q^*$.

Finally, consider H = 4, L = 6. If $C_1^1 = C_h(4)$, then the period 2 sub-game is G(5,3) with $C_h(3) = R_3$. Define q_E4 and q_E5 as the value of q^4 and q^5 when $C_h = R_H$., or:

$$q_E 4(H,L) = \frac{\delta^{2H-1}V\left(1-\delta^{L-H}\right)}{\delta^{L-1}\{W(R_1)-W(R_1-k)\}} \quad \begin{array}{c}\text{when } (L-H) \text{ is even}\\ \text{and } L \text{ is odd}\end{array} \tag{4.3}$$

$$= \frac{\delta^{2H-1}V\left(1-\delta^{L-H-1}\right)}{\delta^{L-1}\{W(R_1)-W(R_1-k)\}} \quad \begin{array}{c}\text{when } (L-H) \text{ is odd}\\ \text{and } L \text{ is odd}\end{array}$$

$$q_E 5(H,L) = \frac{\delta^{2H-1}V\left(1-\delta^{L-H}\right)}{\delta^{L-2}\{W(R_2)-W(R_2-k)\}} \quad \begin{array}{c}\text{when } (L-H) \text{ is even}\\ \text{and } L \text{ is even}\end{array} \tag{4.4}$$

$$= \frac{\delta^{2H-1}V\left(1-\delta^{L-H-1}\right)}{\delta^{L-2}\{W(R_2)-W(R_2-k)\}} \quad \begin{array}{c}\text{when } (L-H) \text{ is odd}\\ \text{and } L \text{ is even}\end{array}$$

Applying the previous result to this sub-game, (HS) must be the unique equilibrium path for $p \leq q_E 4(3,5)$ and $p > q^*$. Now consider the whole game. Provided $q^* > q_E 4(3,5)$, (HS) remains the SPE path for $p \leq q_E 4(3,5)$. When $p \,\varepsilon\, [q_E 4(3,5), q^*]$, the choice is between (DC) and the sequence $\{C_1^1 = C_h$; (DC) from period 2$\}$. Clearly (DC) is preferred by player one since it avoids the period 1 contribution of C_h. For $p > q^*$, the choice is between (DCL) and (HS) and it follows from the definition of q^5 that (DCL) is selected when $p \geq q^5(4,6)$ (this assumes $q^5(4,6) > q^*$).The following proposition generalizes this result.[4]

Proposition 4.1

Suppose $L/2 < H < (L-1)$ where $H > 3$ and the technology is low risk.

(i) Suppose L is odd and $q_E 5\,(H-1, L-1) > q^* > q^4\,(H,L)$. Then the unique SPE path is:

(a) (HS) if $p < q^4\,(H,L)$ or $p \,\varepsilon\, [q^*\; q_E 5\,(H-1,\, L-1)]$,

(b) (DC) if $p \,\varepsilon\, [q^4\,(H,L),\, q^*]$ and

(c) (DCL) if $p > q_E 5\,(H-1,\, L-1)$.

If $p = q^4 (H,L)$, both (HS) and (DC) are SPE paths. If $p = q_E 5 (H-1, L-1)$, both (HS) and (DCL) are SPE paths.

(ii) If L is even, everywhere substitute $q_E 4 (H-1, L-1)$ for $q^4 (H,L)$ and $q^5 (H,L)$ for $q_E 5 (H-1, L-1)$. Then the equilibrium is identical to (i).

Proof See appendix.

The proposition assumes that $q^* > q^4$ which ensures that agents choose to delay the project by (L–H) periods even when the probability p (and so the expected benefit) of the low cost materializing is relatively small. However, it is also assumed that $q^5 > q^*$, which means that the additional disutility from the further one-period delay suffered when players follow (DCL) is relatively large. One implication of these assumptions is that for intermediate probability ranges players may do best by choosing the higher contributions profile (HS) again [i.e., in addition to the choice of (HS) for low probability ranges].

When the first assumption is relaxed and $q^* < q^4$, agents delay the project and follow (DC) only for relatively large values of p. At these large values of p, however, agents are better off switching to the low-funding strategy in period (L–1) and so (DC) is never a SPE path. (HS) is chosen for lower values of p and (DCL) for higher values of p. This case is more likely the larger (L–H) is relative to H. Intuitively, the delay involved in following the path (DC) is large and must be compensated with a high probability p of the low cost being materialized.

Conversely, if the second assumption is relaxed and $q^* > q^5$, then for relatively low values of p agents find it optimal to possibly delay the project (L–H+1) periods and follow (DCL). This implies that the equilibrium where (HS) is chosen for intermediate values of p is eliminated. As p increases, agents choose (HS) then (DC) and finally (DCL). When (L–H) is small relative to H, this scenario is likely. Lemma 4.1 is a special case where $(L-H) = 1$. Here $q_E^4 = 0$ and so (DC) and (DCL) are the only equilibrium paths. Corollary 4.1 summarizes these remarks.

Corollary 4.1

Suppose $L/2 < H < (L-1)$, where $H > 3$ and the technology is low risk. Suppose L is odd. (When L is even, read $q_E4\,(H-1, L-1)$ for $q^4\,(H,L)$ and $q^5\,(H-1)$ for $q_E5\,(H-1, L-1)$].

(i) If $q_E5\,(H-1, L-1) > q^*$ and $q^4\,(H,L) > q^*$, then the unique SPE path is (a) (HS) if $p < q_E5\,(H-1, L-1)$, (b) (DCL) if $p > q_E5\,(H-1, L-1)$ and (c) both (HS) and (DCL) at $p = q_E5\,(H-1, L-1)$.

(ii) If $q^4\,(H,L) > q^* > q_E5\,(H-1, L-1)$, then (HS) is the unique SPE path for $p \leq q^*$ and (DCL) for $p > q^*$.

(iii) If $q_E5\,(H-1, L-1) > q^*$ and $q^4\,(H,L) > q^*$, then the unique SPE path is (a) (HS) if $p < q^4\,(H,L)$, (b) (DC) if $p \in [q^4\,(H,L), q^*]$, (c) (DCL) if $p > q^*$ and (d) both (HS) and(DC) when $p = q^4\,(H,L)$.

The corollary and the discussion prior to it suggest that when (L–H) is large relative to H, both q^4 and q^5 is large relative to q^*.In other words, the disutility involved in delaying the project is significant relative to the expected benefits from "saving" in period L or (L–1). Then for sufficiently large (L–H) relative to H, (HS) should be the unique equilibrium. The next proposition formalizes this intuition.

Proposition 4.2

If $H \leq L/2$, then (HS) is the unique equilibrium path. The project is completed in H periods.

Proof

Suppose (L–H) is even. For $H \leq L/2$, $q_E4(H,L) \geq \{V(1-\delta^{L-H}) / [W(R_1) - W(R_1-k)]\} > 1$ because $V(1-\delta^{L-H}) > V(1-\delta) = W(R_1)$ since $(L-H) > 1$. Similarly, $q_E5(H,L) \geq \{\delta V(1-\delta^{L-H}) / [W(R_2) - W(R_2-k)]\} > 1$ since $\delta V(1-\delta^{L-H}) \geq \delta V(1-\delta^2) = W(R_2)$. Finally, $q^4\,(H,L) > q_E4(H,L)$ and $q^5\,(H,L) > q_E5(H,L)$. Therefore (DC) and (DCL) yield higher payoffs than (HS) only if $p > 1$. When (L–H) is odd, replace $V(1-\delta^{L-H})$ with $V(1-\delta^{L-H-1}) \geq W(R_1)$. This completes the proof.

In fact, a stronger version of Proposition 4.2 holds, as shown below.

Corollary 4.2

Suppose players had the option that if $\tilde{K} = K_\ell$, the project could be completed in period L and no contributions would be required at any time. If $H \leq L/2$, then (HS) is still the unique equilibrium path.

Proof

First, from Proposition 4.2, (HS) constitutes an equilibrium strategy for $H \leq L/2$.

Finally, by completing the project in H periods, player one gets at least $\delta^{2H-1}V$. Finishing the project in L periods costlessly yields $\delta^{L-1}V$. Clearly, $\delta^{2H-1}V \geq \delta^{L-1}V$ for $H \leq L/2$.

Section 5. Risky Technology

In this section, I consider what happens when the cost technology is risky in the sense that k, the difference between the costs, is large. In particular $K_l < R_1$ so that the project can be completed in one period if costs were low for sure (although players must still wait L periods to realize its benefits).

Suppose $H \geq L$, where H is as defined in (3.1). Let L = 2. If $C = C_h$ and $\tilde{K} = K_\ell$, then $M_\ell(H)$, the amount of project costs remaining to be financed in period 2, satisfies:

$$-R_H \leq M_\ell(H) < R_1 \tag{5.1}$$

$M_\ell(H)$ can lie in one of these two sub-intervals:

(a) $0 \leq M_\ell(H) < R_1$

(b) $-R_H \leq M_\ell(H) < 0$

Similarly, if $c_1^1 = K_\ell$ and $\tilde{K} = K_h$, then k is the amount remaining in period 2 and k satisfies:

$$\sum_{q=2}^{H-1} R_q < k \leq \sum_{q=1}^{H} R_q \tag{5.2}$$

k can lie in one of (H−1) possible sub-intervals indexed by l:

(c) $\sum_{q=0}^{l} R_q < k \le \sum_{q-0}^{l+1} R_q, \quad l = 1, 2, \ldots, H-1$

Note that conditions (a) and (b) also imply certain restrictions on k and these must be consistent with the set of restrictions in (c). This requirement leads to these two broad distictions:

Case A : High funding includes conditions (b) and $l = (H-1)$, both of which imply the following restriction on k:

$$\sum_{q=1}^{H-1} R_q < k \le \sum_{q=1}^{H} R_q \tag{5.4}$$

When (5.4) is satisfied $C_h > K_\ell$ so that the period 1 funding is larger when players follow the high contributions profile (HS) rather than the low contributions profile (LS).

Case B: Low funding occurs under conditions (a) and $l < (H-1)$. Here $C_h \le K_\ell$ so that, when players anticipate costs to be high, in period 1 they contribute less relative to when they expect costs to be low.

In solving the model, I will assume that the high-funding case holds. This is consistent with the low-risk technology model of section 3 and 4, where the relevant period 1 contributions were C_h and C_ℓ, and $C_h > C_\ell$ from (3.3).

Define the following three contributions sequences:

(HOS) $C_1^1 = C_h$; from period 2, $\{R_{H-1}, R_{H-2}, \ldots, R_1\}$ if $\tilde{K} = K_h$ or {o, o, ..., o} if $\tilde{K} = K_\ell$.

(LOS) $C_1^1 = C_\ell$; then follow (HOS) from period 2: $\{C_k, R_{H-1}, \ldots, R_1\}$ if $\tilde{K} = K_H$ or $\{0, 0, \ldots, 0\}$ if $\tilde{K} = K_\ell$.

(OOS) $C_1^1 = 0$; then follow (HOS) from period 2: $\{C_h, R_{H-1}, \ldots, R_1\}$ if $\tilde{K} = K_h$ or $\{K_\ell, 0, \ldots, 0\}$ if $\tilde{K} = K_\ell$

Also define :

$$P1(H) = \frac{W(R_H) - W(C_h)}{W(R_H)} \tag{5.5}$$

The next lemma solves all sub-games where L = 2.

Lemma 5.1

(a) If H = 1, then the unique SPE involves completing the game in one period by contributing K_h.

(b) Let L = 2 and $H \geq L$.

(i) If $p < P1(H)$, then the unique equilibrium path is HOS.

(ii) If $p > P1(H)$, then the unique equilibrium path is OOS.

(iii) If $p = P1(H)$, both (i) and (ii) are SPE.

Proof

First, it can be shown that C_h, K_ℓ and 0 are the only period 1 contributions that can occur in any SPE. For example, if $C_1^1 = x$, where $0 < x < K_\ell$, then from period 2 the equilibrium path is $\{(C_h - x), R_{H-1}, \ldots, R_1\}$ or $\{(K_\ell - x), 0, \ldots, 0\}$; so $C_1^1 = 0$ is strictly preferred. Second, player one strictly prefers contributing 0 to K_ℓ in period 1 since from period 3 the paths are identical. Third, comparing (HOS) and (OOS), player one's gain from the latter sequence is a saving of $W(C_h)$ in period 1 but at the cost of a one-period delay relative to period H if $\tilde{K} = K_h$ if in period 2. From Remark 1.2, the latter cost is precisely $(1-p).W(R_H)$ so that cost and gain are equal at p=PI(H). This proves the lemma.

The intuition from Lemma 5.1 suggests that contributing K_ℓ prior to period (L-1) will never be part of an SPE. Since only K_ℓ is required to complete the project in period L and the benefits accrue only in that period, waiting should be the best policy. Similarly, contributing C_h prior to period (L-1) is not a sub-game perfect strategy for any player since in period (L– 1) is not a sub-game perfect strategy for any player since in period (L– 1), the contribution must be zero.[5] The next proposition formalizes this argument.

Proposition 5.1

Suppose $H \geq L$ and the technology is risky. If $K_\ell \geq R_{H+l}$,then in the uniqe SPE players contribute 0 for the first (L-2) periods. From period L–1, the equilibrium path is HOS if $p \leq P1\ (H)$ and 005 if $p > P1\ (H)$.

Proof :

Suppose the proposition is true for L, where L is even and $H \geq L$. Then the SPE path is 0 for (L–2) periods and then for $p \leq P1\ (H)$ { C_h, R_{H-1}, ..., R_1} if $\tilde{K} = K_h$ in period L or { C_h, 0, 0, ..., 0 } if $\tilde{K} = K_\ell$. If $p > P1\ (H)$, the path is {0, C_h, R_{H-1}, ..., R_1} if $\tilde{K} = K_h$ or {0, K_ℓ, 0, ..., 0} if $\tilde{K} = K_\ell$.

Now consider $T = L + l$ and $N = H + l$. $N \geq T$ and T is odd so that player one moves in period L. The candidate profiles are, by the induction hypothesis :

(Ha) {C_h; 0 for (T–3) periods; R_H in period (T–1); and from period T (R_{H-l}, R_{H-2}, ..., R_1) if $\tilde{K} = K_h$ or 0 if $\tilde{K} = K_\ell$}

(Hb) {C_h ; 0 for (T–2) periods ; (R_H, R_{H-1}, ..., R_1) or 0}

(La) {C_ℓ ; 0 for (T–3) periods ; c_k ; (R_H, R_{H-1}, ..., R_1) or 0}

(Lb) {C_ℓ ; 0 for (T–2) periods ; (C_k, R_H, ..., R_1) or 0}

(Oa) {0 for (T–2) periods ; C_h ; (R_H, R_{H-1}, ..., R_1) or 0}

(Ob) {0 for (T–1) periods ; (C_h, R_{H-1}, . ., R_1) or (K_ℓ , 0)}

First, consider the sub-game period T– 1. If $C_1^1 = C_h$,(Hb) is the unique path because 0 is the unique choice in period (T–1). The gain from contributing 0 rather than C_h is W (R_H) and the loss (due to delay) only (1– p). W (R_H). If $C_1^1 = 0$, player two switches from C_h to 0 in period (T– 1) if p > PI (H). If $C_1^1 = K_\ell$, the swith-over probability is :

$$P2(H) = \frac{W(R_H) - W(C_k)}{W(R_H)} \qquad (5.6)$$

Note that P2 > PI for all H since $C_k < C_h$

(i) consider $p \leq P1$ (H). By inspection the path (Oa) is strictly preferred to (Hb) — the unique path under H. The path (Oa) is also strictly preferred to (La) by the same token. This proves the result for $p \leq P1$ (H).

(ii) Suppose $p \varepsilon$ (P1 P2). By inspection, (La) is strictly preferred to (Hb). It remains to be shown that (Ob) yields higher payoffs relative to (La). Comparing the two paths, player one's gain from following (Ob) is a lower contribution level of $\{W(K_\ell) - p.\delta^{L-1} W(K_\ell)\}$. The cost is a one–period delay relative to (La) which is $(1-p).P.\delta^{L-1} W(C_h)$.[6] So, the probability level at which player one is indifferent between (Ob) and (La) is :

$$P3\ (H,L) = \frac{\delta^{L-1}.W(C_h) - W(K_\ell)}{\delta^{L-1}[W(C_h) - W(K_\ell)]} \tag{5.7}$$

If $p < P3$, then (La) is preferred. I will show that $P1 > P3$ and, since $p > P1$, (La) never occurs in any SPE. Showing $P1 \geq P3$ is equivalent to proving that :

$$\frac{W(K_\ell)\ (1-\delta^{L-1})}{W(C_h) - W(K_\ell)} > \frac{\delta^{L-1}.W(C_h)}{W(R_H)}.. \tag{5.8}$$

$$\text{Now the LHS} \geq \frac{W(K_\ell)}{W(C_h)} \geq \frac{W(K_\ell)}{W(R_H)} \geq \frac{\delta^2.W(C_h)}{W(R_H)} > \frac{\delta^{L-1}.W(C_h)}{W(R_H)}$$

I have used the fact that $R_H \geq C_h$ and the assumption that $K_\ell \geq R_{H+1}$.

(iii) for $p \geq p2(h)$, I only need to show that (Ob) is preferred to (Lb). It is sufficient to show that (La) yields higher payoffs relative to (Lb). (Lb) involves a one-period delay relative to (La) with disutility equal to $(1-p).\ \delta^{L-1}W(R_H)$. On the other hand, the savings in contributions is only $(1-p).\ \delta^{L-1}[W(R_H) - W(C_h)]$. This proves the proposition.

For $H < L$, the result is a straightforward generalization of Proposition 4.1. For valuable projects, the (HS) strategy of completing in H periods is still beneficial. The alternative (DCO) here is to contribute 0 for (L–1) periods and then following the AP path.[7] Define:

$$P4(H,L) = \frac{\delta^{2H-3}V(1-\delta^{L-1}) - W(C_h)\ (1-\delta^{L-1})}{\delta^{L-1}[W(C_h)-W(K_\ell) + V(1-\delta^{2H-3})]} \qquad (5.9)$$

P4(H,L) has the following interpretation. Suppose H = L–1. If p = P4(H,L), then player one is indifferent between following (HS) and (DCO) because the benefit from early completion (equal to $\delta^{2H-3}V(1-\delta^{L-1}) - W(C_h)\ (1-\delta^{L-1})$ is equal the expected gains foregone from lower contributions if $\tilde{K} = K_\ell$ in period L [equal to $p.\delta^{L-1}V(1-\delta^{2H-3}) + W(C_h) - W(K_\ell)]$. Let P* = P4(2,3) and PE* be P* evaluated at $C_h(2) = R_2$ and $K_\ell = 0$. Then PE* = $W(R_2/[W(R_1) + W(R_2)]$.

Proposition 5.2 Suppose H < L and $K_\ell \geq R_{H+1}$.[8]

(i) If p < PE*, (HS) is the unique SPE.

(ii) If p > PE*, (DCO) is the unique SPE.

(iii) If p = PE*, both (HS) and (DCO) are SPE.

Proof

First consider the case H = 2, L = 3. Applying Lemma 5.1, it is easy to see that the following are the candidate SPE paths:

(HS) $\{C_h, R_1\}$

(DCO) $\{0, 0, (C_h \text{ or } K_\ell), (R_1 \text{ or } 0)\}$ if $p \geq P1(2)$.

(DC) $\{0, C_h, (R_1 \text{ or } 0)\}$ if $p \leq P1\ (2)$

(La) $\{K_\ell, C_k, (R_1 \text{ or } 0)\}$ for $p \leq P2(2)$.

(Lb) $\{K_\ell, 0, (C_k \text{ or } 0), (R_1 \text{ or } 0)\}$ for $p \geq P2(2)$.

Note that P2(2) > P1(2) by definition. Suppose p ≤ P2(2). By observation, player one prefers (DC) to (La). Evaluating player one's utility from (DC) at the point p = P1(2) yields $V_1^1 = [W(R_2) - W(C_h)] \cdot \frac{\delta^2 W(R_1)}{W(R_2)} + \delta^3 V$. Since $V_1 < [\delta V - W(C_h)]$, players one's utility from (HS), (HS) is the unique equilibrium path for p ≤ P1(2).

Next, consider $p > P1(2)$. Clearly, the path (DCO) is strictly preferred to the path (Lb). (DCO) also yields higher expected utility relative to (La). To show this, note that $U_1^1(DCO) = [(1-P)(\delta^3 V - \delta^2 W(C_h)) + p(\delta^2 V - \delta^2 W(K_\ell))] \geq [(1-p)(\delta^3 V - W(K_\ell)) + p(\delta^2 V - \delta^2 W(K_\ell))] > [(1-p)\delta^3 V + p\delta^2 V - W(K_\ell)] = U(La)$. The first inequality follows because $K_\ell \geq R_{H-1}$ and $C_h \geq R_h$.

Comparing (DCO) and (HS), it is easy to check that player one prefers (HS) when $p < P^*$ and (DCO) when $p > P^*$. Further, $P^* > P1(2)$ because:

$$P^* = \frac{W(R_2) - W(C_h)(1-\delta^2)}{\delta^2 [W(R_1) + W(C_h) - W(K_\ell)} > \frac{W(R_2) - W(C_h)}{\delta^2 [W(R_1) - W(K_\ell)} > P1(2)$$

This completes the proof of the proposition.

The generalization to any $H = L-1$ follows immediately. Consider $H = 3$, $L = 4$. The candidate SPE paths are (HS) if $p \leq PE^*$, (Hb): $\{C_h, 0, 0, (R_2 \text{ or } 0), (R_1 \text{ or } 0)\}$ if $p \geq PE^*$, (DC): $\{0, C_h, R_2, (R_1 \text{ or } 0)\}$ if $p \leq P4(3,3)$. and (DCO): $\{0, 0, 0, (C_h \text{ or } K_l), (R_2 \text{ or } 0), (R_1 \text{ or } 0)\}$ if $p \geq P4(3,3)$. The result follows because

$$P4(3,3) = \frac{W(R_2) - W(C_h)\{(1-\delta^2)/\delta^2\}}{W(C_h) - W(K_\ell) + W(R_1) + W(R_2)} \text{ and } PE^* > P4(3,3)$$

since $C_h > K_\ell$ from (5.4). In general, $PE^* > P(H,L)$ for $H > 2$ and $L > H$ and so the proposition is true for any $H = L-1$. Further, this argument is valid for any $H \leq (L-1)$ since PE^* is independent of H and L. This completes the proof.

Section 6. Characteristics of the Equilibrium

In this section, I show that (under an additional assumption) the project may not be completed whenever there is some proability (no matter how small) that the project would be unfinished in a full information world.

Specifically, suppose $K_h > \sum_{q=0}^{\infty} R_q$ and $\sum_{0}^{S-1} R_1 < K_\ell \leq \sum_{0}^{S} R_q$.

Then the following lemma is proved in the appendix:

Lemma 6.1

Suppose $L = 2$ and $S > 1$ Define $e1\ (s) = W(C_\ell\ ) / W(R_s)$. Then:

(i) If $p > e1(S)$, the unique SPE is $C_1^1 = C_\ell$, and from period 2, (LS) if $\tilde{K} = K_\ell$ or 0 if $\tilde{K} = K_h$. The project is completed in S periods only if $\tilde{K} = K_\ell$. Otherwise the project is not completed.

(ii) If $p < e1\ (S)$, then the unique SPE path is $C_1^1 = 0$ and, from period 2, (LS) if $\tilde{K} = K_\ell$ or 0 if $\tilde{K} = K_h$. The project is completed in (S+1) periods if $\tilde{K} = K_\ell$ and is not finished if $\tilde{K} = K_h$.

(iii) If $p = e1(S)$, both (i) and (ii) are SPE.

Note that any equilibrium path which involves a positive contribution in period 1 entails a risk since the project may not be completed with probability p. In fact, for p sufficiently low, the SPE posits a zero contribution in period 1.

Now suppose there is a minimum contribution of $x > 0$ that is required in period 1. The minimum is "small" in that $W(x) \leq \delta^2 . W(C_\ell\ )$. x can be interpreted as the amount of investment required to "learn" $\tilde{K}$ in period L. Define $e2(S) = W(x) / \delta^{2S-1}V$.

Proposition 6.1

If $p < e2(S)$ and $W(x) \leq \delta^2 . W(C_\ell)$, then the project will never be completed.

Proof

Define $e3(S) = [W(C_\ell) - W(x)] / W(R_S)$. From Lemma 6.1, it follows that:

(a) If $p \geq e3\ (S)$, the SPE is identical to part (i) of Lemma 6.1.

(b) If $p \leq e3\ (S)$, the SPE path is $C_1^1 = x$ and, from period 2, 0 if $\tilde{K} = K_h$ or $\{c_\ell - x,\ R_{S-1},\ ...,\ R_1\}$ if $\tilde{K} = K_\ell$.

Next, observe that $e2\ (S) \leq e3\ (S)$ because $W(R_S).W(x) \leq$

$\delta^{2S-1}V[W(c_f) - W(x)]$ [use the definition $W(R_S) = \delta^{2S-3}(1-\delta^2)$ and the assumption $W(x) \le \delta^2.W(c_f)$].

Finally, if $p \le e3\ (S)$ and player one follows the SPE path in part (ii), his expected utility is $p\delta^{2S-1}V - W(x)$. So when $p < e2\ (S) \le e3\ (S)$, his expected utility from participating in the project is negative. Since he can always get at least zero by not participating, the result follows.

Essentially, under the asumptions of Proposition 6.1, player one's individual rationality constraint is violated. Strictly speaking, this is not an inefficiency relative to the "first best" world[9] since a social planner would also not participate under the same circumstances. The following two results, proved by example in the appendix, identifies the types of inefficiencies that could arise in the bargaining procedure.

Proposition 6.2

Suppose H < L and the technology is low risk. If (HS) is the SPE path, agents' total contributions will exceed the true project cost if the true costs are low. A social planner's total contributions may or may not exceed the true project cost even when they are low.

Proposition 6.3

Suppose the technology is risky. If (DCO) is the SPE path, then agents contribute 0 for (L–1) periods and complete the project in (L–1+H) periods if the true costs are high. This time until completion may or may not be longer than the corresponding time in the first-best optimum. It will always be longer than the corresponding time in the full-information optimum.

(6.2) suggests that the bargaining procedure may be inefficient in that agents may over-fund the project (ex-post this will appear like a cost overrun). (6.3) says that the inefficiency may involve an "unnecessary" delay in project completion (in addition to that in a full-information world).

Section 7. Conclusion

I considered the problem of how two agents would provide for a joint project when the cost of the project is intially uncertain.

Agents alternate in making contributions to the project. These contributions should be considered as sunk costs which are borne by the agent in the period in which the payment is made. Initially, agents have access to a high cost technology but, in some future period L, a cheaper technology may become available.

Suppose the cost technology is low risk in the sense that the difference between the high and low cost is "small." If the parameters of the model are such that, in any equilibrium, the project cannot be completed prior to the revelation period L then two kinds of agreements are possible. One, prior to period L agents contribute as if costs were high for sure and then follow along the full-information Admati-Perry equilibrium path. Two, prior to period (L–1), agents again follow the high contributions schedule (HS) but in period L–1 switch to the low contributions schedule (i.e., they contribute as if costs were low for sure. Note that, in period L–1, the lower cost technology is not yet available). If agents' anticipation is proved to be wrong and the low-cost technology does not materialize in period L, they re-switch to (HS).

If the parameters of the model are such that the project can be completed prior to period L by implementing the high-cost technology, then the possible equilibrium paths are:

(HS) completing the project prior to period L. Here players' total contributions may exceed the true project cost if costs are revealed to be low in period L. This is likely to occur if L is large relative to the high cost.

(DC) Delay completion by making no contributions initially; then follows (HS) to finish in L periods.

(DCL) Follow (DC) up to period (L–1) and then switch to the low contributions schedule. Both (DC) and (DCL) occur if L is small relative to the high cost.

The second set of results relate to the case of risky technology i.e., when the cost differences are "large". For parameter values such that project completion post–dates L in equilibrium, agents contribute nothing prior to period L. For other parameter values, early completion via (HS) is again the unique equilibrium when L is large relative to the high cost.

Two types of inefficiencies (relative to the first–best optimum) were identified. First, agents may contribute in excess of the true cost–thus over-funding the project. Second, agents may delay the project "unnecessarily" by contributing nothing for several periods. This delay is in addition to that caused by agents lack of commitment in the full–information world.

The results of this paper is sensitive to the particular assumption regarding how the two costs affect outcome. Without this assumption, the low contributions schedule would always be chosen. An alternative is to make the cost difference a random variable–this would allow the result to be generalized to technologies of any degree of risk.

APPENDIX

Proof of Lemma 4.1 First I prove a result for H=2

Claim 1 Suppose H=1 and L=3. If $C_h(2)=R_2$, then the SPE path is (a) (DC) if $p \leq q1\,(2)$ and (b) (HS) if $p \geq q1\,(2)$.

Proof The candidate profiles are :

(HS) $\{C_h(2), R_1\}$

(DC) $\{0, C_h(2), (R_1 \text{ or } R_1 - K)\}$ if $p \leq q1\,(2)$

(DCL) $\{0, C_\ell(2), (k \text{ or } R_1), (R_1 \text{ or } 0)\}$ if $p \geq q1\,(2)$

(LS) $\{C_\ell(2), k, (R_1 \text{ or } R_1 - k)\}$ if $p \leq q3\,(2)$

(LLS) $\{C_\ell(2), 0, (k \text{ or } R_1), (R_1 \text{ or } 0)\}$ if $p \geq q3\,(2)$.

Observe that $q3 \leq q1$. Consider $p \leq q1\,(2)$. Clearly (DC) is preferred to (LS) by player one. Comparing (HS) and (DC), $q4\,(2,3) = 0$ at $C_h(2) = R_2$. When $p \geq q3\,(2)$, (DCL) is preferred to (LLS). Comparing (HS) and (DCL), player one's utility under (DCL) is strictly lower than $\delta^3 V$ and therefore his utility under (HS). When

$p \,\varepsilon \left[q1\ (2)\ q3\ (2)\right]$, compare (DCL) to (LS). For $p > W\{C_\ell(2)\}/\delta^2\{W(R_1)-W(R_1-k)\} = q6$ player one prefers (DCL) to (LS). Now the result for H=3 is unaffected by whether q6 $\gtrless$ q3. So, without loss of generality, assume that q6 > q3. This proves the claim.

Now consider H=3 and L=4.

Claim 2 Suppose H=3 L=4 and $C_h\ (3) < R_3$. Then :

(i) If $q^* > q5\ (3,\ 4)$, the unique SPE path is (DC) if $p \leq q^*$ and (DCL) if $p \geq q^*$.

(ii) If $q^* < q5\ (3,\ 4)$, the unique SPE path is (DC) if $p \leq q^*$, (HS) if $p \,\varepsilon \left[q^*,\ q5\ (3,\ 4)\right]$ and (DCL) if $p \geq q5\ (3,\ 4)$.

Proof The candidate profiles are :

(HS) $\{C_h\ (3),\ R_2,\ R_1\}$ if $p \geq q_E1\ (2)$

(HOS) $\{C_h\ (3),\ 0,\ R_2,\ (R_1 \text{ or } R_1 - k)\}$ if $p \leq q_E1\ (2)$

(DC) $\{0,\ C_h\ (3),\ R_2,\ (R_1 \text{ or } R_1 - k)\}$ if $p \leq q^*$

(DCL) $\{0,\ C_h\ (3),\ R_2, -k,\ (k \text{ or } R_1),\ (R_1 \text{ or } 0)\}$ if $p \geq q^*$

The paths with $C_1^1 = C_\ell(3)$ are ignored since (DC) and (DCL) will always be preferred to them. Also observe that $q_E1\ (2) = q^*$. So consider $p \leq q^*$. By observation, (DC) is always preferred to (HOS). If $p \geq q^*$, (HS) is preferred to (DCL) if $p \leq q5\ (3,4)$ by definition of q5. The result follows immediately.

If $C_h(3) = R_3$, note that $q_E5\ (3,4) = 0$ so that $q^* > q_E5\ (3,4)$. This proves part (a) of the lemma. Now consider H = 4, L = 5. If $C_1^1 = C_h(4)$, in the period 2 sub-game G(4,3) $C_h(3) = R_3$ and so player two's optimal contribution is 0. This rules out the path (HS) for H = 3 and so for all H > 3. This proves part (b) of the lemma.

Proof of Proposition 4.1

Suppose the result holds for $H = N$ and $L = T$, where T is even and $T/2 < N < (T-1)$. Consider $H = N+1$ and $L = T+1$. By the induction hypothesis, the candidate SPE profiles in the G(L,H) game are (HS) if $p \leq q_E 4(N-1, T-1)$ and $p \, \varepsilon \, [q^*, q_E 5 (n,T)]$; (DC) if $p \varepsilon [q^4 (N-1, T-1), q^*]$ and (DCL) if $p \geq q5 (N,T)$. Profiles involving zero contributions at any time other than period 1 are ignored since they can never be part of any SPE path.

Suppose $p \leq q^*$ and compare (HS) and (DC). Since L is odd, q4(H,L) is defined. Further $q4 \; (H,L) = q4 \; (N+1, T+1)$

$$= q_E 4(N-1, T-1) - \frac{W(C_h) \; (1-\delta^{T-N})}{W(R_1) - W(R_1-k)} < q_E 4(N-1, T-1) \, .$$

Then, by definition of q4, the SPE path is (HS) for $p \leq q4(H,L)$ and (DC) for $p \; \varepsilon \, [q4(H,L), q^*]$.

Suppose $p > q^*$ and compare (HS) and (DCL). Since $L = (T+1)$ is odd, player one does not move in period $(L-1)$ and so does not benefit from the switch to (LS) in that period. Therefore, player one's utility is strictly higher under (HS) and so (HS) is the SPE path for $p \, \varepsilon \, (q^*, q_E 5(N,T))$, the probability range over which (HS) is defined. This proves part (i) of the proposition. An analogous proof holds for part (ii).

Proof of Lemma 6.1 The candidate profiles are :

$$\text{(LOS)} \; \left\{ C_\ell ; \left(R_{H-1}, \ldots, R_1 \right) \text{ if } \tilde{K} = K_\ell \text{ or } 0 \text{ if } \tilde{K} = K_h \right\}$$

$$\text{(O)} \; \left\{ 0, (C_\ell, R_{H-1}, \ldots, R_1) \text{ or } 0 \right\}$$

$$\text{(HL)} \; \left\{ C_h ; \{(R_{H-1} - k), R_{h-2}, \ldots, 0\} \text{ or } 0 \right\}$$

Clearly the choice is between (LOS) and (0). The period 1 benefit is $W(C_\ell)$ if (o) is followed but at the cost of a one-period delay relative to period S if $\tilde{K} = K_\ell$. Since this cost is $p.W(R_s)$, the result follows.

Proof of Proposition 6.2

First, I provide an example where $L = 2$ but both the bargaining and planning procedures result in immediate completion. The data assumed is: $\delta = 0.5$, V = 10, W(.7) = 0.09, W(1) = 3.75, W(1.2) = 5, W(2.1) = 10^{10} and W(.) is linear between these points. These numbers imply $R_1 = 1.2$ and $R_2 = 1$.

Example 1 : No over-investment

Assume $K_h = 1.2$, $K_\ell = 0.7$ and L = 2. Check that k = 0.5 < R_1 as required. Since H = 1, the bargaining solution is immediate completion by contributing 1.2. If the planner does the same, his utility is V(HS) = 20 = W(1.2) = 15. If he wants to wait till period 2 his utility is strictly is less than δ(2V) = 10.

Example 2 : Over-investment in bargaining

Assume $K_h = 2.19$, $K_\ell = 1.4$, L = 3. Since H = 2, by Claim 1 of the appendix, (HS) will be a SPE path for some probability values, where:

(HS) $\{(K_h - R_1), R_1\} = \{0.99, 1.2\}$

If the planner also wants early completion, (HS) is his optimal path with utility V(HS) = (0.5)20 − W(.99) − (0.5) W(1.2) = 3.872. But there is a better path:

(EQ) {0.7, 0.7, (0.7 or 0)}

Check that V(EQ) = $(.5)^2$ 20 − W(.7) (1+0.5) − (1−p). $(0.5)^2$ W(0.7) = 5 − (0.09 + 0.045) − (1−p) (0.0225) ≥ 4.8425 > V(HS).

Proof of Proposition 6.3

Assume the same data used in the previous proof (other than the cost data) and follow the same procedure.

Example 1: No delay

Suppose $K_h = 2.9$, $K_\ell = 0.69$ and L = 2. Then H = 3,

$K_\ell < R_1$, $K > (R_1 + R_2)$ and $K_\ell > R_4 = (0.25)^2(3.75) = 0.015$ as required. I will show that both bargaining and planning involves completion in four period if $\tilde{K} = K_h$ and two periods if $\tilde{K} = K_\ell$. In bargaining, (DCO) is the unique SPE path for p > 0.9 where:

(DCO) {0, (.7 or .69), (1 or 0), (1.2 or 0)}.

The planner's best path is:

(EQ) {.7, (.7 or 0), (.7 or 0), (.8 or 0)}

and yields utility V(EQ) = 2.18(1−p) + 10p

If he wants to finish earlier, his best option is {.7, (1 or 0), (1.2 or 0)} with utility equal to [1.79 (1−p) + 10p] < V(EQ).

Example 2: Delay In Bargaining

With the same data, asume that L = 4. In bargaining, (DCO) is a SPE path where:

(DCO) {0, 0, (0 or .69), (1 or 0), (1.2 or 0)}

Project completion is in six periods if $\tilde{K} = K_h$ and four periods if $\tilde{K} = K_\ell$. The planner's best path is:

(EQ) {.7, .7, .7, (.8 or 0)}

where V(EQ) = 2.18. He finishes in exactly four periods. Observe that under full-information completion is in three periods if $\tilde{K} = K_h$ and one period if $\tilde{K} = K_\ell$, which is earlier than the first-best.

NOTES

1. The disutility costs of contributing 0 instead of R_1 in period 1 are (1) a one-period delay in project completion time and (2) changes in the sequence of contributions. By definition, R_q is equal to the sum of these costs. For $q \geq 2$, both these costs are decreasing in q Between q = 1 and q = 2, however, although the first cost still decreases the second cost increases which explains the ambiguity of the relationship between R_1 and R_2.

2. This calculation follows directly from Remark 1 2. Consider the sub-game beginning in period (L−H+1). Then player one is the second mover in that sub-game and gets $\delta^{2H-2}V$ in equilibrium, or $\delta^{L-H}.\delta^{2H-2}V$ in the whole game

3. Strictly speaking, since $C_h < R_h$ and H = 3, the statement of Lemma 4.1 does not cover this case. However, the result is similar and the proof in the appendix does cover H = 3 as well as H = 2.

4. To keep the statement of the proposition simple, I will ignore the non-generic equilibrium path where there is a zero contribution in period 1 and then agents follow (HS).

5 This is because the delay cost in contribution 0 rather than C_h is $(1-p)W(R_H)$ while benefits are exactly $W(R_H)$.

6. This can be computed as follows Suppose the game started in period L and player one contributed R_H instead of C_h in that period From Remark 1.2, the disutility from delay is (1–p) W(R_H) Adjusting for the C_h contribution and discounting, the period 1 disutility works out as given

7. Note that, by Lemma 5.1, an equilibrium path will never involve a contribution of C_h prior to period (L–1) The path {0, ., 0; C_h in period (L–1); . } yields strictly lower utility relative to (HS) for p< P1(H)

8. For H = 2, L = 3 substitute P* for PE* everywhere

9. In a "first best" world, contributions are made by a single agent with cost function W(.) and benefit 2V from project completion

REFERENCES

1. A R. Admati and M. Perry, "Joint Projects Without Commitment", Working Paper, Stanford University and Harvard University, 1988.

2. C. Bliss and B. Nalekuff, Dragonslaying and ballroom dancing: The private supply of a public good, *Journal of Public Economics 25* (1984), 1–12.

3 P. Bolton and J. Farrel, "Decentralization, Duplication and Delay," Working Paper, Harvard University, Berkeley and Hoover, 1988.

INTERNATIONAL TRADE

2

On Keynesian Unemployment in a Fix-Flex Price Model of a Small Open Economy With Two Variable Inputs*

Bhaswar Moitra

1. Introduction

Recent studies of factor unemployment in monetary models of small open economies emphasize the non-Walrasian features of such models. This class of models essentially extends the single factor, single output framework of the Barro and Grossman (2,3) and Malinvaud (13) fix-price closed economy macroeconomic models to the case of small open economies operating under a fixed exchange rate regime. Dixit (8) in a seminal paper argues that it is important to provide a systematic choice - theoretic foundation to the standard analysis of monetary phenomena in non-Walrasian models of open economies. Unfortunately, his single commodity and small economy assumptions restrict the scope of his analysis to a particular type of unemployment regime where the wage rate, fixed above the Walrasian market clearing level, constrains suppliers in the labor market while all domestic agents are unconstrained in all other markets. Neary (16) extends the model by disaggregating the production sector into a non-traded and a traded good sector. This allows him to generate all the non-Walrasian equilibrium regimes described by Barro and Grossman and Malinvaud within a unified framework.[1] As noted earlier, such models have two common features[2] : a) all prices are fixed and quantities adjust to clear markets; and b) there is a single variable input. The first feature distinguishes these models from the neo-classical (Walrasian) models of balance of payment where all prices adjust infinitely fast [see Takayama (19) and Dixit and Norman (9)]. These features also distinguish them from the Heckscher-Ohlin type models of unemployment developed by Brecher (6) and Helpman (11) where some prices are fixed, there are two variable and sectorally mobile inputs (capital and labour) and the rental rate

---

* I would like to thank Rajat Deb for suggesting that I explore this topic and Josef Hadar for some very helpful comments. I would like to thank Jonathon Burke for several conversations about the Arrow-Debreu model. The remaining errors are, of course, my own,

adjusts to clear the capital market. The Brecher and Helpman models differ in some respects. Brecher considers two traded goods whose prices are internationally fixed and domestic agents are unconstrained at these prices. Helpman analyzes a model with a non-traded and a traded good. The price of the traded good is exogenously determined and domestic agents face no quantity constraints at these prices. The price of the non-traded good adjusts to clear the market. Further, the real wage rate is fixed in terms of an index of commodity prices, implying that the money wage rate is variable.

These models, however, have some important common features, First, some prices are fixed while others adjust infinitely fast. Second, labor unemployment arises purely because of a high real wage rate. Third, domestic agents face no quantity constraints in the product markets. Fourth, there are two variable, sectorally mobile inputs that are substitutes in each production process. This substitutability may actually exacerbate the degree of labour unemployment although this is not clear. Finally, it is not at all clear why the economy will produce both goods. In fact, Brecher's model is Richardian as pointed out by Bhagwati and Srinivasan (5).

All these raise some interesting questions. First, while there is an enormous body of literature on general equilibrium models where all prices either adjust at infinite speed or are completely fixed, not many systematic studies of the mixed case seem to exist. Benassy (4) studies the case where monopolistic agents adjust prices at the beginning of each period, but prices remain fixed within the period. Negishi (17) generates stationary prices in the presence of underemployment as a strategic equilibrium under specific expectational assumptions imposed on agents. The question what behavioral and expectational assumptions we impose on agents if some prices are fixed a priori and others adjust at infinite speed to generate mutually consistent actions remains. We shall address this question for a specific set of circumstances.

Second, we shall extend the Heckscher-Ohlin type models to consider the case of Keynesian unemployment, i.e., unemployment generated by demand failures. What we shall find is that such demand failures are possible even if some prices adjust at infinite speed. This is scarcely novel, but models of this type with a consistent choice theoretic basis are scarce. Finally, we shall ensure that our equilibria are characterized by positive output levels for all final goods.

We conclude this section by briefly reviewing the contents of the rest of the paper. We consider a small, open, monetary economy producing and consuming two goods: a non-traded good and a traded good. There are two freely mobile factors of production whose total stocks are historically given. The analysis is essentially intertemporal though our interest is in temporary or period equilibria. Within this structure we shall generate Keynesian unemployment equilibria. We shall assume, for simplicity, that the money supply is completely sterilized within each period. Adjustments in the economy's money stock via changes in international reserves occur at the end of each period. In sections 2 — 4 we describe the basic features of our model. In section 5 we formally define what we mean by Keynesian equilibria in our model and prove its existence. In section 6 we study the effect of changes in government demand for the non-traded good. Section 7 contains some concluding remarks.

2. Preliminaries

We consider a small, open, perfectly competitive, private ownership economy operating under a fixed exchange rate regime. The economy is small in the usual sense: domestic agents do not face quantity constraints in the traded goods markets at the fixed exogenously determined (world) price for these goods. The economy is perfectly competitive in the sense that all agents treat prices as parameters in their decision process.

In the rest of this section we describe the nature of the goods and their prices; the types of agents, ownership and endowments; and provide an intuitive description of the notion of equilibrium.

Time

The time horizon is visualized as the real line. This is partitioned into disjoint, right-open intervals of finite length called market periods. For any market period $[r, r+1) \subseteq R$ the point r is called the beginning of the period and the intervals $[-\infty, r)$ and $[r+1, \infty]$ are called the past and the future, respectively. Our concern is with the behavior of the economy within an arbitrary market period and not with its evolution across market periods. It is in this sense that our model describes short run phenomena as opposed to medium or long term movements.

Goods and Prices

There are five types of goods in the system : two produced goods, a non-traded good (good n) and a traded good (good t), both produced and consumed domestically; two factors of production, designated labor and capital, whose stocks at the beginning of each period are historically given; and a good that serves as the sole unit of account, medium of exchange and store of value, henceforth called money. The stock of money at the beginning of each period is exogenously determined.

The quantities of the goods are represented by the non-negative real numbers X_n, X_t, L, K and M respectively. The prices of the non-monetary goods in terms of units of money are represented by the non-negative real numbers p_n,p_t,w (wage rate) and r (rental rate) respectively. The price of money in terms of itself is, by definition, equal to one. We assume that within the period the price of the non-traded good and the nominal wage rate remain fixed. The price of the traded good is exogenously determined. Assuming that this is fixed during the current period, at a fixed exchange rate, it is also fixed in terms of domestic money. The rental rate is assumed to be completely flexible.

Agents

There are three types of agents : a single household; two firms, one in each sector; and the government. The household owns all productive resources, i.e., the factor inputs and the firms. We assume that the firms are independently managed. The government controls the stock of money in the economy, the exchange rate, and buys final goods which it finances through (for simplicity) money printing.

Equilibrium

We focus on a specific type of equilibria : Keynesian temporary equilibria. While we refrain from presenting a formal definition till section 5, it is useful to specify the general features of our equilibrium notion at this juncture. Our equilibria are characterized by two features.

a) individual actions (defined more carefully in the next section) are compatible in the *current* period; and b) at least one firm faces demand constraints in its output market and the household faces demand constraints in at least one input market.

Remark 1

This is a general description of the type of equilibria we are interested in. In our model, by assumption, domestic agents are unconstrained in the market for the traded good. Hence, at equilibrium, only the firm producing the non-traded good faces quantity constraints in the output market. Further, since the rental rate is assumed to be completely flexible and adjusts to clear the capital good market, the household can only face quantity constraints in the labour market.

3. Microtheoretic Foundations

In this section we develop the micro-theoretic foundations of our model. We begin by briefly reviewing some concepts and specify some notational conventions we shall adopt throughout the paper. These will be used to describe the behaviour of agents and to characterize the equilibria of the economy.

As is by now well known, the principal distinguishing feature of non-Walrasian general equilibrium models is that at equilibrium all agents may not be able to transact at their *notional* (i.e., Walrasian) levels. In general, some agents in each market may face quantity constraints on their desired trades. Such quantity constraints are assumed to be determined by pre-specified rationing rules. An agent's *effective* demand (supply) in a market is his demand for (resp. supply of) the good taking into account his quantity constraints in all *other* markets under the prevailing price system.[3] If no quantity constraints are perceived in other markets, effective singnals coincide with the notional ones. An agent's *actual* trade is the quantity he is able to transact at equilibrium. It may be noted that for each agent notional, effective and actual trades coincide under a Walrasian equilibrium price system.

Before proceeding further we clarify some additional behavioural assumptions of our model. There are five markets in the system : the markets for the two final goods; the input markets; and the money market. Since, all prices are measured in terms of the domestic currency, a price system for the economy is given by the vector $(p_n, p_t, w, r) >> 0$. In addition, we assumed that p_n, p_t, and w are fixed within the current period and r adjusts to clear the market for capital. The model then is a blend of the Walrasian flex-price and the non-Walrasian (pure) fix-price models. The underlying behavioral assumptions about trading possibiliities perceived by agents in these two types of models are radically different. Given *any* price system, Walrasian agents do not perceive any quantity constraints on their net trades and, hence, emit *notional* demand and supply signals. On the other hand, given a price system, agents on the long side of the market in a non-Walrasian fix-price model, anticipate quantity constraints and, hence, emit *effective* demand and supply signals by taking these into account. What we need is a similar behavioural assumption for our model. An appropriate one is a blend of the ones described above. In particular, we assume that our agents do not perceive any quantity constraints in the markets for money, capital and the traded good and take into account quantity constraints in all other markets.

We adopt the following notational conventions. An agent's notional demand for (supply of) a good will be designated by a superscript, d, (resp s), viz. X_i^d, X_i^s, for notional household demand for and notional domestic supply of good i, respectively. Effective demands and supplies will be designated by an additional symbol tilde (~) over the relevant variable, viz. $\tilde{X}_i^d$. Perceived quantity constraints are designated by a circumflex (∧) above the relevant variable, viz., $\hat{X}_i$ Finally, all prices and quantities that are exogenously determined and treated as datum in the current period will be designated by a bar (—) over the relevant variable. viz. $\bar{M}, \bar{L}, \bar{K}$ which represent the historically determined stocks of money, labour and capital at the beginning of the period, respectively.

We now proceed to describe the behaviour of the agents in our system.

Production

In each final good sector the firm maximizes profits in the current period. There are no inventories. The sectoral production functions are given by

$$X_n = F_n (K_n, L_n) \tag{1}$$

$$X_t = F_t (K_t, L_t) \tag{2}$$

in the non-traded and traded goods sectors, respectively, where L_i and K_i represent the quantities of inputs used in the production of good i, i = n,t. We make the following assumptions on technology.

Assumption 1

The function $F_i : R_+^2 \rightarrow R_+$ are homogenous of degree α_t, i = n,t, twice continuously differentiable and strictly concave.

Assumption 1 implies that equations (1) and (2) may be rewritten in intensive form as

$$X_n = L_n^{\alpha} n \, f_n(k_n) \tag{3}$$

$$X_t = L_t^{\alpha} t \, f_t(k_t) \tag{4}$$

$L_i \neq 0$. It follows that : $\frac{\partial F_i}{\partial K_i} = L_i^{\alpha_i - 1} f_i'(k_i)$ and $\frac{\partial F_i}{\partial L_i} = L_i^{\alpha_i - 1}$

$[a_i \; f_i \; (k_i) - k_i \; f_i' \; (k_i)]$, where ki = $\frac{K_i}{L_i}$, for $L_i \neq 0$. It should be observed that for some $\alpha_i \; \varepsilon \; (0,1)$, $L_i^{\alpha_i}$ can take both positive and negative values. For notational simplicity we ignore the absolute value (I.I) and treat $L^{\alpha}i$ as positive for all $\alpha_i > 0$

Assumption 2

a) $f_i' > 0, \; f_i'' < 0, \; k_i^{\mathrm{Lim}} \rightarrow \infty \; L_i^{\alpha_i - 1} \; f_i' = 0, \; k_i^{\mathrm{Lim}} \rightarrow 0 \; L_i^{\alpha_i - 1} \; f_i' = \infty$

b) $\left(\alpha_i f_i - k_i f_i'\right) > 0$,

$$k_i^{\text{Lim}} \to \infty \; L_i^{\alpha_i - 1}\left(\alpha_i f_i - k_i f_i'\right) = \infty, \quad k_i^{\text{Lim}} \to 0 \; L_i^{\alpha_i - 1}\left(\alpha_i f_i - k_i f_i'\right) = 0$$

Assumption 2 imposes curvature restrictions on the F_i (.) functions. In particular, Assumption 2(b) requires that $\frac{\partial F_i}{\partial L_i}$ be unbounded as L_i goes to zero and is zero as L_i goes to infinity. These are critical in the proof of Proposition 2. Finally, we make a simplifying assumption on technology. This makes the proof of Proposition 2 remarkably simpler without harming the essential aspects of the model.

Assumption 3

$\alpha_n = \alpha \in (0,1)$ and

$\alpha_t = 1$

Assumption (3) requires that the technology in the non-traded good sector satisfies diminishing returns to scale and that in the traded goods sector satisfies constant returns to scale.

The profit function in sector i, i=n,t, is given by

$$\Pi_i = \bar{p}_i F_i \,(L_i, K_i) - \bar{w}\, L_i - r\, K_i \tag{5}$$

In the absence of quantity constraints, input demands and, hence, output supplies are determined by maximizing the profit function with respect to the variables L_i and K_i in each sector, yielding the notional trades :

$$L_i^d = L_i^d(\bar{p}_i, \bar{w}, r) \tag{6}$$

$$K_i^d = K_i^d(\bar{p}_i, \bar{w}, r) \tag{7}$$

$$X_i^s = X_i^s(\bar{p}_i, \bar{w}, r) \tag{8}$$

i = n,t. In the presence of quantity constraints on demand, profit maximization reduces to minimization of the cost of production of the (constrained) output level $\hat{X}_i^s$. This yields the effective input demands

$$L_l^d = L_l^d(\bar{p}_l, \bar{w}, r, \hat{x}_l^s) \tag{9}$$

$$K_l^d = K_l^d(\bar{p}_l, \bar{w}, r, \hat{x}_l^s) \tag{10}$$

Note that for the producing the traded good, the effective demands coincide with the notional demands.

Remark 2

In the presence of constant-returns to scale technology, output and input demands in the traded good sector are indeterminate for any given output and factor price vector in a partial equilibrium framework. As we shall show in section 5 (see the proof of Proposition 2) this indeterminacy is resolved in general equilibrium.

Remark 3

Assumption 1 and the assumption that the firm producing the traded good does not face quantity constraints in any market implies that $\Pi_t = 0$. It is also clear that $\Pi_n \geq 0$. This leaves one with the distressing feeling that the firm facing the demand constraint could be making a positive profit while the one maximizing profits without any demand constrains makes zero profits. The only plausible reason for such a phenomenon is that firms do not have equal access to technology. This, of course, is the essential reason why profits may differ in the Arrow-Debreu Model.

The Behaviour of Households

At the beginning of the current period, the household has a stock of maney ($\bar{M}$), labor services ($\bar{L}$) and capital ($\bar{K}$). We assume that factor inputs have no intrinsic consumption value and, hence, are inelastically supplied.

Assumption 4

The initial stock of capital and labour in each period is strictly positive and finite, i.e., $(\bar{L}, \bar{K}) \in (0,\infty) \times (0,\infty)$; and the initial stock of money $\bar{M}$ is finite and non-negative, i.e., $\bar{M} \in (0,\infty)$.

The household makes decisions about input supplies, consumption of final goods and terminal money holdings in the current period given the resource constraints, current and expected prices, quantity constrains and profit income, on the basis of well defined intertemporal preferences. Following Muellbauer and Portes (14) we assume that the household's (intertemporal) preferences can be represented by a real valued function

$$V\ (X_n, X_t, \bar{L} - L, \bar{K} - K, M; \theta\ )$$

where X_i is the consuption of good i, L and K are the supplies of labor and capital respectively, and M the terminal money balance in the current period. It is assumed that the part of the parameters that govern the probabilistic expectation of future prices, profits and quantity constraints that depend on the current period values of X_i, L, K and M have been absorbed in the form of the function $V\ (. ; \theta\ )$ and the vector θ represents the remainder.[4] Presumably, θ depends on current prices, rental rate, wage rate and profit distribution. We make the following assumption about the function $V(. ; \theta\ )$.

Assumption 5

$V\ (. ; \theta\ )$ is strictly concave and monotonically increasing in X_n, X_t and M.

We assume that the current periods profits are distributed in the current period. In our model the household faces quantity constraints only in the labor market. Therefore, its decision problem reduces to solving two optimization exercises;

A. $$\underset{X_n, X_t, M, L, K}{\text{Max}} \quad V\ (X_n, X_t, \bar{L} - L, \bar{K} - K, M; \theta\ )$$

Subject to :

i) $0 \le L \le \bar{L}$

$0 \le K \le \bar{K}$

(resource constraint), and

ii) $$\Pi_n + \bar{M} + \bar{w}L + r\bar{K} \ge \sum_{i=n}^{t} \bar{P}_i X_i + M$$

(budget constraint); and

B. $$\underset{X_n, X_t, L, K, M}{\text{Max}} \quad V\ (X_n, X_t, \bar{L} - L, \bar{K} - K, M; \theta\ )$$

Subject to

i) $0 \le L \le \hat{L}$ (Labour supply constraint)

ii) $0 \le K \le \bar{K}$ (Capital endowment constraint) and

iii) $$\bar{M} + \Pi_n + \bar{w} L + r \bar{K} \ge \sum_{i=n}^{t} \bar{P}_i X_i + M$$

Solution of A yields the notional demand and supply functions :

$$\tilde{X}_i^d = \tilde{X}_i^d\left(\bar{p}_n, \bar{p}_t, \bar{w}, r, \bar{M}, \bar{K}, \bar{L}\ \theta\right) i = n,t \qquad (1,1)$$

$$M^d = M_i^d\left(\bar{p}_n, \bar{p}_t, \bar{w}, r, \bar{M}, \bar{K}, \bar{L}\ \theta\right) \qquad (12)$$

$$L^s = L^s\left(\bar{p}_n, \bar{p}_t, \bar{w}, r, \bar{M}, \bar{K}, \bar{L}\ \theta\right) = \bar{L} \qquad (13)$$

$$K^s = K^s\left(\bar{p}_n, \bar{p}_t, \bar{w}, r, \bar{M}, \bar{K}, \bar{L}\ \theta\right) = \bar{K} \qquad (14)$$

Equations (13) and (14) state that inputs are inelastically supplied.[5] Further, only equation (13) is relevant since the effective demands for the final goods and terminal money balance and effective supply of capital are determined by taking into account the labor supply constraint $\hat{L}$. These are deduced by solving B :

$$\tilde{X}_i^d = \tilde{X}_i^d\left(\bar{p}_n, \bar{p}_t, \tilde{I}, +\bar{M};\ \theta\right) \qquad (15)$$

$$\bar{M}^d = \bar{M}^d\left(\bar{p}_n, \bar{p}_t, \tilde{I}, +\bar{M};\ \theta\right) \qquad (16)$$

$$\tilde{K}^s = \tilde{K} \qquad (17)$$

Where $\tilde{I} = \tilde{\Pi}_n \bar{w} \hat{L} + r\bar{K}$ (18)

It is obvious that for B to be meaninful, $\hat{L} < \bar{L}$.(see equation 13). We make the following assumptions on the properties of equation (15).

Assumption 6

a) $\tilde{X}_n^d$ (.) is continuously differentiable in $\tilde{I}$;

b) $$0 < p_n \frac{\partial \tilde{X}_n^d}{\partial \tilde{I}} < 1$$

c) $p_{n_j} \le p_{n_{j+1}}$, $j=1,2,\ldots$, implies

$$\tilde{X}_n^d\left(p_{n_j}, \bar{p}_t, p_{n_j}\overline{X}_n + \bar{p}_t\overline{X}_t + \overline{M};\theta\right) \ge$$

$$\tilde{X}_n^d\left(p_{n_{j+1}}, \bar{p}_t, p_{n_{j+1}}X_n + \bar{p}_t\overline{X}_t + \overline{M};\theta\right)$$

d) $\lim_{p_n \to \infty} \tilde{X}_n^d\left(p_n, \bar{p}_t, p_n\overline{X}_n + \bar{p}_t\overline{X}_t + \overline{M};\theta\right) = 0$

Assumption 6(a) is self explanatory. Assumption 6(b) requires that for all p_n the marginal propensity to consume the non-traded good out of current income lies between zero and one. The spirit of the assumption is clearly Keynesian. Assumption 6(c) requires that given p_t, x_t, x_n and $\overline{M}$ if p_n increases then the effective demand for the non-traded good falls. Assumption 6(d) extends assumption 6(c) to the limiting case.

Govermnent Behavior

The government buys final goods for its own consumption which it pays for by printing money. Its budget constraint is given by

$$\Delta M = \bar{p}_n G_n + \bar{p}_t G_t \tag{19}$$

where G_i, $i=n$, t is the government's demand for good i in the current period and ΔM is the quantity of money it prints in the current period to finance its deficit. We make the following assumptions on the government's demand for final goods.

Assumption 7

a) $G_n > 0$,

b) $Gn < \left(\frac{\overline{K}}{k_n}\right)^{\alpha} f_n(k_n)$, $0 < k_n < \infty$

Assumption 7(a) is self explanatory. Assumption 7(b) ensures that domestic production of the non-traded good exceeds the government's demand if all of the capital stock is used to produce the non-traded good.

4. Allocation of Rations

Given the single household an single firm (in each market) structure of our model, allocation of rations (quantity constraints) is particularly simple. All we require is that agents on the long side of a market (other than the market for capital) face a quantity constraints equal to the short side quantity signal. This yields:

$$\hat{X}_n = \min\left[X_n^s, \tilde{X}_n^d + G_n\right]$$

$$= \tilde{X}_n^d + \text{ [Assumption 4(b)]} \qquad (20)$$

$$\text{and } \hat{L} = \min\left[L^s, \tilde{L}_n^d + L_n^d\right] < L^s = \bar{L} \qquad (21)$$

for the labor market constraint on the household to be strictly binding. since domestic agents are unconstrained in the traded good market, there are no ration allocations in this market. We assume that agents are unconstrained in the money market. As will become clear in a moment this implies that in the presence of a fixed exchange rate, the government does not impose any foreign exchange restrictions (see Proposition 1). Finally, in the market for capital, agents behave like Walrasian traders, i.e., they do not perceive any quantity constraints under any price system and quantity restrictions in other markets.

The next question is, how do we characterize an equilibrium of the system? Further, does such an equilibrium exist in our model. We deal with these questions in the next section.

4. Keynesian Equilibrium

Given any arbitrary price system and a set of quantity constraints, agents signal their effective demands and supplies in each market. In the event that such signals are not mutually compatible *and* the household faces demand constraints in the labor market while the firm producing the non-traded good faces demand constraints in the output market, equations (20) and (21) are used to update the demand constraints. If, in addition, $\bar{K}_n^d + K_t^d \neq \bar{K}$ the rental rate is also changed according to some predetermined rule. The prices of the two final goods and the wage rate are held fixed throughout the process. An equilibrium of the system is a price system and a set of prices and effective demand signals such that given the price system :

(a) The effective demand signals that generated the quantity constraints remain unchanged after the agents carry out their respective optimization programs using these quantity constraints; and (b) the effective excess demand in the capital market is zero.

At such an equilibrium, actual trades in the non-traded good and labor markets are determined by equations (20) and (21) respectively. Actual trades equal effective demands and/or supplies in all other markets. This leads to the following definition.

Definition (Keynesian Equilibrium)

A Keynesian equilibrium of our economy is an eleven tuple of real numbers (p_n^*, p_t^*, w^*; r^*, X_n^*, X_t^*, K_n^*, K_t^*, L_n^*, L_t^*, M^*) satisfying :

$$p_i^* = \bar{p}_i, \quad i = n,t,$$

$$w^* = \bar{w}$$

prices of the final goods and the wage rate remain fixed;

$$X_i = L_i^{\alpha_i} f_i(k_i), \quad i = n,t, \alpha_n \in (0,1), \quad \alpha_t = 1 \tag{3),(4}$$

production functions;

$$\bar{w} = \frac{\alpha f_n(k_n) - k_n f_n'(k_n)}{f_n'(k_n)} \tag{22}$$

cost minimizing condition in th non-traded good market;

$$\bar{w} = p_t \left[f_t(k_t) - k_t f_t'(k_t) \right] \tag{23}$$

$$r = p_t f_t'(k_t) \tag{24}$$

profit maximizing conditions in the traded good market;

$$k_n \bar{L}_n^d + k_t L_t^d = \bar{k} \tag{25}$$

capital market clearing conditions;

$$X_n = \min \left[\bar{X}_n^d + G_n, X_n^{s^*} \right]$$

$$= \tilde{X}_n^d \left(p_n, p_t, p_n X_n + p_t X_t^s, + \bar{M}; \theta \right) + G_n \tag{26}$$

producers in the non-traded good sector are demand constrained;

and $M=\tilde{M}^d(p_n,p_t,p_nX_n+p_tX_t+\overline{M};\ \theta)$ (27)

households are unconstrained in the money market.

It should be observed that such an equilibrium need not exist for any arbitrary configuration of money and factor stocks, and preassigned prices of the final goods and wage rate. Hence, one has to prove that given the initial stocks of money and capital resources there is a configuration of preassigned final goods prices, wage rate and labor stock for which the economy has an equilibrium satisfying (a) and (b) above.

Proposition 1

At equilibrium.

$$\tilde{M}^d-\Delta M-\overline{M}=p_t(X_t^s-\tilde{X}_t^d-G_t)$$

This is a well known result in the literature on open economy macro-economics : the excess of demand for terminal money balances over the current periods money supply equals the value of the surplus in the current account in domestic currency. The proposition implicitly requires that there are no quantity restrictions imposed on foreign exchange.

We now proceed to state our main result.

Proposition 2

Given Assumptions 1—7, there exists an equilibrium for our economy if $\overline{L}$ is sufficiently large.

Remark 4

It is clear from the proof of Proposition 2 (please see Appendix) that given $(\overline{p}_t,\overline{w},\overline{M},\overline{G}_n,\overline{K})\rangle\rangle 0$ and $0\langle\alpha\langle 1$. There is a range of values of p_n for which a Keynesian equilibrium exists for the economy. Further, any change in the values of these parameters will lead to a change in the critical range of values of p_n. Hence, for any given valve of $\overline{p}_n$ within range for a given set of value of α, $\overline{p}_n,\overline{w},\overline{M}$ and $\overline{G}_n$ a change in value of any of these parameters may lead to a switch in equilibrium regimes. It follows, therefore, that any comparative static exercise is necessarily restricted to changes in parameter values that do not cause regime switching for a given value of p_n.

6. Comparative Statics : Change in Government demand for the non-traded good.

We look at one comparative static property of our model : the effect of a change in the level of government demand for the non-traded good. Given a set of values of $\bar{p}_t, \bar{p}_n, \bar{w}$, and $\bar{M}$ we assume that the change in the value of is G_n sufficiently small so that the new equilibrium is also Keynesian (see Remark 4). The homogeneity property of the production functions considerably simplifies the analysis of the stability properties of the model for changes in G_n. In particular, a change in G_n has no effect on and k^*_i, i = n, t and r^*. Given the continuity properties of the capital market equilibrium condition and the production functions the problem reduces to ensuring that the demand for labor in the non-traded good sector converges to the new equilibrium value determined by equation (28). This turns out to be rather straightforward, so we omit the proof.[6] The actual computation of the effect of a change in G_n on the endogenous variables then becomes extremely simple. This we think is a very attractive feature of our model.

Differentiation of Equations (31), (4), (25), and (26) yields

$$\begin{bmatrix} 1 & 0 & -\alpha f_n \tilde{L}_n^{\alpha-1} & 0 \\ 0 & 1 & 0 & -f_t \\ 0 & 0 & k_n & k_t \\ 1-\bar{p}_n \dfrac{\delta \tilde{X}_n^d}{\delta \bar{I}} & -\bar{p}_t \dfrac{\delta \tilde{X}_n^d}{\delta \bar{I}} & 0 & 0 \end{bmatrix} \begin{bmatrix} dX_n \\ dX_t \\ d\bar{L}_n \\ dL_t^d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ dG_n \end{bmatrix}$$

We drop the (*) symbols for notational convenience.

Solution of equation (29) yields :

$$\frac{dX_n}{dG_n} = \frac{\alpha k_t f_t(k_t) L_n^{\alpha-1}}{A} \tag{30}$$

$$\frac{dX_t}{dG_n} = \frac{k_n f_t}{A} \tag{31}$$

$$\frac{d\tilde{L}_t}{dG_n} = \frac{-k_t}{A} \tag{32}$$

and

$$\frac{dL_t^d}{dG_n} = \frac{k_n}{A} \tag{33}$$

where $A = -\overline{p}_t k_n f_t'(k_t)\frac{\delta \tilde{X}_n^d}{\delta \tilde{I}} - \left(1 - \overline{p}_n \frac{\delta X_n^d}{\delta \tilde{I}}\right) \delta k_t f_n(k_n)\, \tilde{L}_n^{\alpha-1}$. Noting that $\frac{\delta \tilde{X}_n^d}{\delta \tilde{I}}$ is positive and $\overline{p}_n \frac{\delta \tilde{X}_n^d}{\delta \tilde{I}} \in (0,1)$ [assumption 7(b)] we conclude that A<0 and further that $\frac{dX_n}{dG_n}$ and $\frac{d\tilde{L}_n^d}{dG_n}$ are positive, and $\frac{dX_t}{dG_n}$ and $\frac{d\tilde{L}_t^d}{dG_t}$ are both negative. It also follows from equations (32) and (33) that $\frac{dL}{dG_n} = \frac{d\tilde{L}_n^d}{dG_n} + \frac{dL_t^d}{dG_t} \lesseqgtr 0$ according as $K_n \lesseqgtr k_t$. This allows us to go one step further. Remembering that $\tilde{I} = \tilde{\Pi}_n + \overline{w}\, L + r\, \overline{K}$, and that the profit in the non-traded good sector increases with output, we can conclude that $\frac{d\tilde{I}}{dG_n} > 0$ if $K_n > k_t$. Hence if the traded good is normal, the current account balance $B = \overline{p}_t(X_t^d - \tilde{X}_t^d - G_t)$ declines (rises) if G_n rises (resp. falls). The picture is not clear if $K_n \leq k_t$.

7. Concluding Remarks

In this paper we studied Keynesian unemployment in a small dependent economy with two sectorally mobile inputs. The model developed may be viewed as an attempt to amalgamate recent work on non-Walrasian equlibrium models of International Trade with the Heckscher-Ohlin type models of unemployment in open economies. The salient features of our model are : (1) there are two variable inputs only one of whose prices is fixed a priori ; (2) the production functions satisfy homogeneity properties; and (3) the equilibrium notion is one of temporary Keynesian equilibrium.

The first two features are similar to those found in the literature on Heckscher-Ohlin type models of underemployment in open economies. These features, particularly the first, distingush our model from the open economy macroeconmic models of recent vintage. The third feature extends the Heckscher-Ohlin type models to the case of Keynesian unemployment.

In sections 2 - 4 we developed the microeconomic foundations of the model. This was necessary because carefully developed general equilibrium models with both fixed and flexible prices are scarce in the literature. Inspite of the fact that models with such features are abundant in the literature scant attention appears to have been paid to the agents' perceptions of trading possibilities under this scenario. Once one observes that assumptions made about these behavioural specifications of the system differ radically in the flex price (Walrasian) and the fix-price (non-Walrasian) general equilibrium models, development of a counterpart in a model with both fixed and flexible prices assumes critical importance. We have tried to provide such a behavioural specification.

In section 5 we defined Keynesian equilibnum and proved that such equlibria exist in our model. Given the homogeneity properties of the production functions our equilibrium specifications bear a close resemblance to those popularized by Kemp (1964) in the context of Walrasian models of open economies. The homogeneity properties also simplify the comparative static analysis considerably although one has to be extremely careful about this (see Remark 4.)

The proof of Theorem 2 highlights an important feature of models of this type: alternative types of equilibria are consistent only with particular configurations of the parameters of the system. This, of course, is expected.

The model may be extended in several directions. What we have tried to do is to develop a logically consistent microeconomic foundation for a general equlibrium model of trade when some prices are fixed and others adjust at ifinite speed under a particular set of institutional and equlibrium specifications. Similar models can be developed for different types of disaggregation of the production sector and dfferent equilibrium specifications. The general methodology may also serve to provide proper microfoundations to models with flexible exchange rates. What is more interesting, perhaps, is a general model where the parameter space may be clearly divided into subsets in accordance with their compatibility with alternative types of equilibria. This is the subject of the author's current research.

REFERENCES

1. Apostol, T.M. "Mathematical Analysis," Addison-Wesley Publishing Company, Reading, mass., 1974.

2. Barro, R J. and Grossman, H I., A General Disequilibrium Model of Income and Employment, American Economic Review, 61 (1971), 82-93.

3. Barro, R.J. and Grossman, H.I., "Money, Employment and Inflation," Cambridge University Press, Cambridge, 1976

4 Benassy, J.P., "The Economics of Market Disequilibrium," Academic Press, New York.

5. Bhagawati, J.N. and Srinivasan, T.N., "Lectures on International Trade," M.I.T. Press, Cambridge, Mass., 1983.

6. Brecher, R., Optimal Commercial Policy for a Minimum Wage Economy, J. Int., Econ 4, (1974), 139-50.

7. Cuddington, J.T., Johansson P.O., and Lofgren, K.G., "Disequilibrium Macroeconomics in open Economies," Basil Blackwell, Oxford, 1984.

8. Dixit, A.K., The Balance of Trade in a Model of Temporary Equilibrium with Rationing, Rev. Econ. Stud. 45, (1978), 393-404.

9. Dixit, A.K. and Norman, V., "The Theory of International Trade : A Dual General Equilibrium Approach," Cambridge University Press, London, 1980.

10. Drazen, A., Recent Developments in Macroeconomic Disequilibrium Theory, Econometrica 48, (1980), 283-306.

11. Helpman, E , Macroeconomic Policy in a Model of International Trade with a Wage Restriction, Int. Econ. Rev. 17, (1974), 262-77

12. Kemp, M.C , "*The Pure Theory of International Trade,*" Prentice-Hall, Englewood-Cliffs, N.J., 1964

13. Malinvaud, E , "*The Theory of Unemployment Reconsidered,*" Basil Blackwell, Oxford, 1977

14. Muellbauer, J and Portez, R., Macroeconomic Models with Quantity Rationing, *Econ. J.* 88, (1978), 788-821.

15. Neary, J P Nontraded Goods and the Balance of Trade in Neo-Keynesian Temporary Equilibrium, Quart. *J. Econ.* 95, (1980), 403-29.

16. Neary, J P., and Stiglitz, J E , Towards a Reconstruction of Keynesian Economics . Expectations and Constrained Equilibrium, *Quart J. Econ.* (*supplement*) 98, (1983), 199-228.

17. Negishi, T , "*Microfoundations of Keynesian Macroeconomics,*" North-Holland Publishing Co., Amsterdam, 1979.

18. Silvestre, J., fix-Price Analysis in Exchange Economies, *JET,* 26, (1982), 28-58.

19. Takayama, A , "*International Trade,*" Holt, Reinhart and Winston, Inc., New York, 1972.

Endnotes

1. Neary's model is unsuitable for analyzing commercial policy questions where a strict distinction between exportables and importables is required. For an extension along these lines see Cuddington et al. (7) and the studies cited there.

2. This is not strictly true. Neary (15) and Cuddington et al. (1984) briefly mention the case of Orthodox Keynesian unemployment where the price of the non-traded good adjusts to clear the market. Cuddington et al. also consider the case of flexible exchange rates. In both these cases, however, agents' perceptions about trading possibities are not clearly described.

3. We adopt the Clower-Benassy definition. For alternative definitions see Drazen (10) and Silvestre (18).

4. See Muelbauer and Portes (14) p 798-800 and also Benassy (4) chapter 8 and Appendix K for a discussion about the logic of this formulation.

5. Neary and Stiglitz (16) leave out the input supplies as arguments of the direct-indirect utility function.

6. Let $\dot{L}_n - \beta h\,(L_n, \bar{p}_n, G_n)$, $0 < \beta < 1$.

APPENDIX

Proof of Proposition 1

Summation of the budget constraints of the agents in our model yields :

$$\bar{w}(\tilde{L}_n^d + L_t^d) + r(\tilde{K}_n^d + K_t^d) + \bar{p}_n(\tilde{X}_n^d + G_n) + \bar{p}_t(\tilde{X}_t^d + G_t) + \tilde{M}^d$$

$$-\bar{p}_n X_n + \bar{p}_t X_t^s + \Delta M + \bar{M} + \bar{w}L + r\bar{K}$$

Noting that at equilibrium $X_n = \tilde{X}_n^d + G_n$, $L = \tilde{L}_n^d + L_t^d$ and $K = \tilde{K}_n^d + K_t^d$ we conclude that :

$$\tilde{M}^d - \Delta M - \bar{M} = \bar{p}_t(X_t^s - \tilde{X}_t^d - G_t) .$$

Proof of Proposition 2

(I) Let $w^* = \bar{w}$ and $p_t^* = \bar{p}_t$ be finite and positive. The Assumption 1 and 2(b) imply that there exists a finite $k_t^* > 0$ satisfying

$$\bar{w} = \bar{p}_t\left[f_t(k_t) - k_t f_t'(k_t)\right] .$$

Given k_t^*, and assumptions 1 and 2(a), there exists a positive, finite r^* satisfying

$$r^* = \bar{p}_t f_t'(k_t^*) .$$

Hence, assumptions 1 and 2 imply that there exists a positive and finite k_n^* satisfying

$$\frac{\bar{w}}{r^*} = \frac{[\alpha f_n(k_n) - k_n f_n'(k_n)]}{f_n'(k_n)} .$$ Also, k_n^*, k_t^* and r^* are unique.

Noting that $k_i = k_i^*$, $i = n, t$, we demonstrate first that there exists a p_n^* for which there are $\tilde{L}_n^d > 0$, $L_t^d > 0$ satisfying equations (3), (4), (25) and (26). Hence, letting $\bar{L} > \tilde{L}_n^{d*} + L_t^{d*}$ we conclude the proof of the theorem.

(II) From equations (3), (4), (25) and (26) we get on straightforward maniplulation :

$$\tilde{X}_n^d(p_n, p_t, p_n L_n^\alpha f_n^* + \frac{\bar{K}}{k_t^*} L_n \bar{p}_t f_t^* - \frac{k_n^*}{k_t^*} L_n \bar{p}_t f_t^* + \bar{M}, \theta) + G_n - L_n^\alpha f_n^* = 0 \qquad (28)$$

where $f_i^* = f_i(k_i^*)$, $i = n, t$. Since, k_i^{α} is defined for $L_i \neq 0$, the left hand side of equation (28) is defined for $L_n \in (0, \frac{\overline{K}}{k_n^*})$. Define :

$$g(L_n, p_n) = \tilde{X}_n^d[p_n, \overline{p}_t, p_n L_n^{\alpha} f_t^* + \overline{p}_t f_t^* (\frac{\overline{K}}{k_t^*} - \frac{k_n^*}{k_t^*} L_n) + \overline{M}; \theta] + G_n - L_n^{\alpha} f_n^*,$$

$$L_n \in (0, \frac{\overline{K}}{k_n^*}),\ p_n \in (0, \infty).$$

$$\lim_{L_n \to 0} g(L_n, p_n) = \tilde{X}_n^d(p_n, \overline{p}_t, \frac{\overline{K}}{k_t^*} \overline{p}_t f_t^* + \overline{M}; \theta) + G_n$$

and

$$\lim_{L_n \to \frac{\overline{K}}{k_n^*}} g(L_n, p_n) = \tilde{X}_n^d\left[p_n, \overline{p}_t, p_n f_n^* (\frac{\overline{K}}{k_n^*})^{\alpha} + \overline{M}; \theta\right] + G_n - (\frac{\overline{K}}{k_n^*})^{\alpha} f_n^*$$

Define :

$$h(L_n, p_n) = \begin{cases} \lim_{L_n \to 0} g(L_n, p_n), & L_n = 0 \\ g(L_n, p_n), & L_n \,\varepsilon\, (0, \frac{\overline{K}}{k_n^*}) \\ \lim_{L_n \to \frac{\overline{K}}{k_n^*}} g(L_n, p_n), & L_n = \frac{\overline{K}}{k_n^*} \end{cases}$$

h (Ln, pn) is continuous on the interval $[0, \frac{\overline{K}}{k_n^*}]$. Further, $h(0, p_n) > 0$ [assumption 7(a)] and $h(\frac{\overline{K}}{k_n^*}, p_n) \lesseqgtr 0$. Since g(.) is differentiable, h(.) is differentiable in the interior of its domain, for each p_n.

Hence, $\frac{\partial}{\partial L_n} h\,(L_n, p_n) = \alpha L_n^{\alpha-1} f_n^*\left(\frac{\partial X_n^d}{\partial \tilde{I}} - 1\right) - \frac{\partial X_n^d}{\partial \tilde{I}}\, \bar{p}_t\, f_t^*\, \frac{k_n^*}{k_t^*} < 0$

$L_n \varepsilon\ (0, \frac{\bar{K}}{k_n^*})$. [from (I)′ and assumption 8 (b)]

Since, $\forall L_n \in \left[0, \frac{\bar{K}}{k_n^*}\right]$ and $\forall\, j = 1,2,\ldots\ldots\ p_{n_j} \leq p_{n_{j+1}}$ implies

$$\tilde{X}_n^d\left[p_{n_j}, \bar{p}_t, p_{n_j} L_n^\alpha f_n^* + \left(\frac{\bar{K}}{k_t^*} - \frac{k_n^*}{k_t^*} L_n\right)\bar{p}_t\, f_t^* + \bar{M};\ \theta\right]$$

$$\geq \tilde{X}_n^d\left[p_{n_{j+1}}, \bar{p}_t, p_{n_{j+1}} L_n^\alpha f_n^* + \left(\frac{\bar{K}}{k_t^*} - \frac{k_n^*}{k_t^*} L_n\right)\bar{p}_t\, f_t^* + \bar{M};\ \theta\right],$$

$h\,(L_n, p_{n_j}) \geq h(L_n, p_{n_{j+1}})$, $\forall_j = 1, 2, \ldots$ and $L_n \in \left[0, \frac{\bar{K}}{k_n^*}\right]$.

Further,

$$\lim_{p_n \to \infty} \tilde{X}_n^d\left[p_n, \bar{p}_t, p_n L_n^\alpha f_n^* + \left(\frac{\bar{K}}{k_t^*} - \frac{k_n^*}{k_t^*} L_n\right)\bar{p}_t\, f_t^* + \bar{M};\ \theta\right]$$

$= 0,\ \forall L_n \in \left[0, \frac{\bar{K}}{k_n^*}\right]$, given k_n^*, k_n^t, implies

$$\lim_{p_n \to \infty} h(L_n, p_n) \overset{\text{pointwise}}{=} G_n - L_n^\alpha f_n^* = \bar{h}(L_n)$$

on $L_n \in \left[0, \frac{\bar{K}}{k_n^*}\right]$. Note that $\bar{h}(L_n)$ is continuous and monotonically declining in L_n on $\left[0, \frac{\overline{K}}{k_n^*}\right]$. Also, $\overline{h}(0) - G_n > 0$ and $\overline{h}(\frac{\overline{K}}{k_n^*}) - G_n - (\frac{\overline{K}}{k_n^*})^\alpha f_n^*$ which is less than zero if $G_n < (\frac{\overline{K}}{k_n^*})^\alpha f_n^*$.

That is the limit function $\overline{h}(L_n)$ as $p_n \to \infty$ is continuous, monotnically declining and assumes both positive and negative values on its domain.

Summarizing : (i) The functions $h(l_n, p_n)$ are continuous and monotonically declining on the (compact) interval $\left[0, \frac{\overline{K}}{k_n^*}\right]$ for each p_n;

(ii) $\lim_{p_n \to \infty} h(L_n, p_n) \overset{\text{(pointwise)}}{=} \overline{h}(L_n)$, which is continuous, monotonically declining, and assumes both positive and negative values;

(iii) $p_{n_j} < p_{n_{j+1}} \Rightarrow h\left(L_n, p_{n_j}\right) \geq h\left(L_n, p_{n_{j+1}}\right) \forall L_n \in \left[0, \frac{\overline{K}}{k_n^*}\right]$.

Hence, by Dini's theorem [Apostol (1)], $h\left(L_n, p_{n_j}\right)$ converges uniformly to $\overline{h}(L_n)$ for any increasing sequence of prices $\left(p_{n_j}\right)$, i.e., for each $\varepsilon > 0$, $\exists$ J such that $\forall j > J$, $\left| h\left(L_n, p_{n_j}\right) - \overline{h}(L_n) \right| < \varepsilon$.

Pick $0 < \in^* < \left(\frac{\overline{K}}{k_n^*}\right)^\alpha f_n^* - G_n$. Then, $\exists J^*$ such that $\forall j > J^*$, $\left| h\left(L_n, p_{n_j}\right) - \overline{h}(L_n) \right| < \left(\frac{\overline{K}}{k_n^*}\right)^\alpha f_n^* - G_n$, i.e., $h\left(\frac{\overline{K}}{k_n^*}, p_{n_j}\right) < 0$.

We know that for all p_n, $h(0, p_n) > 0$ and further, $\frac{\partial}{\partial L_n} h\left(L_n, p_{n_j}\right) > 0$.

Hence, for all $p_{n_j} > p_{n_{j^*}}$, $\exists L_n^* \in \left(0, \frac{\bar{K}}{k_n^*}\right)$ such that $h\left(L_n, p_{n_j}\right) = 0$. Also, L_n is unique for each $p_{n_j} > p_{n_{j^*}}$. That is, $\exists p_n$ such that equation (28) is satisfied. To see that this p_n is finite, let $\left(p_{n_j}\right) = (j)$. Pick $p_n^* = \bar{p}_n > j^*$. Let L_n^* be the solution to $h(L_n, \bar{p}_n) = 0$. Then $L_t^* = \frac{\bar{K}}{k_t^*} - \frac{\bar{K}}{k_t^*} L_n^* > 0$. Also, $K_i^* = L_i^* K_i^* > 0$, $i = n, t$.

From equation (3) $X_n^* = (L_n^*)^{\alpha} f_n^*$; and from equation (4) $X_t^* = L_t^* f_t^*$.

Finally, $M^* = \tilde{M}^d(\bar{p}_n, \bar{p}_t, \bar{p}_n X_n^* + \bar{p}_t X_t^* + \bar{M}; \theta)$.

This concludes our proof.

International Capital Movement, Trade Policy and Unemployment

Sugata Marjit and Rajat Acharya

Introduction

The purpose of this paper is to reconsider the role of certain conventional policies in increasing employment in an open economy. In particular we shall focus on taxation of foreign capital, import tariff and export promotional measures in a typical less developed economy and find their impacts on unemployment. Recently Batra and Lahiri (1987) developed a model of intersectorally mobile capital and unemployment [as a contribution to the literature popularised by Corden and Findlay (1975)], They consider welfare ranking of alternative trade policies and conclude that export promotional schemes undertaken by the LDCs should prove to be beneficial to them.

In this paper we consider a north-south trade model where north represents the advanced region and south indicates the group of LDCs lumped into one economy. The south has a fixed money wage rate and suffers from unemployment whereas labor force is fully employed in the north. We have an integrated production structure where the south not only imports capital from the north but also purchases intermediate input produced in the north. Actually one can think of a situation where the south assembles imported intermediaries with its labor and sells the final goods in the rest of the world. This is essentially trade with specialisation in the stages of production. Such structures have been extensively discussed by Marjit (1987, 1988). Efficacy of policies pursued in the south will be shown to depend on the production structure of the north. A tax on foreign capital may be beneficial in terms of improving the employment situation while tariff may fail to generate larger employment. In the literature on tariff and unemployment [see Chan (1982), Sauern-heimer (1986)] labor intensities of contracting and expanding sectors are compared to find out the net impact of tariff. We shall show that even if the contracting sector in the south commands a larger share in total employment to start with, the expansionary effect following an increase in capital-tax or a decline in tariff, may lead to larger employment.

Unlike the models in the literature, we emphasize on international capital mobility and structure of trade to discuss the policy impacts. The paper will be divided into three sections. In the first section we describe the determination of equilibrium. In the second section we put forward the basic propositions. In the last section we conclude the paper after suggesting some extensions and generalisations.

Section–1 : The Model

We assume two countries north and south. North produces three traded goods — exportable good Y^*, importables x^* and an intermediate product I. This intermediate good is used in production of good x^* in combination with labour. South produces only exportables good-x and importables good-y. But it lacks the technology to produce the intermediate good. Thus south imports good-I and combines it with unskilled labour to produce its export good. South also has no endowment of capital. Northern capital uses southern labour, both skilled and unskilled, to produce southern import good y. Commodity prices in these economies are assumed to be fixed.[1] This is a simplifying assumption which helps us concentrate on the supply side of the adjusted proces only. Price of intermediate is, however, determined endogenously. Northern capital and labour, as well as southern skilled workers are fully employed. In both countries wage differentials exists due to heterogeniety of labour, though such heterogeniety is asymmetric. Unlike in south, northern workers are assumed to be "specialised" : both L^*_1 and L^*_2 are skilled, but job-specific, i.e., while L^*_2 can be used in production of both northern good-x^* and good-I, it cannot be used to produce Y^*. In a sense, norther workers are mobile across stages of production but immobile across sectors producing different final goods. This type of job specificity is not uncommon to find in advanced industrialised nations. Finally, southern unskilled workers are unemployed and can be hired at a fixed money wage rate assumed to be a markup over subsistence.

---

[1] We can justify fixity of final good prices by assuming very high value of elasticity of substitution in demand. In the extreme case of infinite elasticity prices are fixed. For every high but less than infinite elasticity, price changes will be very small and supply effects will dominate demand effects. Thus we can treat prices variable allowing for low values of elasticity of substitution in demand, and our model will work for high values of that elasticity.

To describe the model formally, we shall use the following symbols. Variables with asterisk will denote northern variables.

p_j, p^*_j = price of j-th good (j = x, y, I)

W = fixed money wage rate of southern unskilled workers.

W_s = money wage rate of southern skilled workers

W^*_1 = money wage rate of northern labour used in production of y^*.

W^*_2 = money wage rate of northern labour used in production of x^* and I.

r = return to capital

L_e = total employment of southern unskilled labour

L^*_1, L^*_2 = given supply of northern workers

K^* = stock of northern capital

S = given supply of southern skilled workers

a_{ij} (i = L, K, S, I ; j = X.Y) = input coefficient in south

a^*_{ij} (i = L, K, I ; j = x,y,I, i ≠ j) = input coefficient in north

We assume that production functions obey constant returns to scale and diminishing marginal returns to imputs. Markets are competitive. For simplicity we assume that unskilled labour is required in fixed proportion, though foreign capital and skilled southern workers are substitutes of each other.[2] Usually, highly unskilled labour can only be poor substitute for more sophisticated foreign inputs. Similarly, we assume that northern intermediary sector operates under fixed input coefficients. This is indeed a simplifying assumption. Then competitive equilibrum implies,

$$P_I a_{Ix} + W a_{Lx} = P_x \tag{1}$$

---

[2] We could have assumed substitution elasticity between skilled and unskilled workers but that would only complicate the analysis and conditions underlying employment improvement will only be altered marginally.

$$r\, a_{ky} + W_s\, a_{sy} + W\, a_{Ly} = P_y \tag{2}$$

$$r\, a^*_{ky} + w^*_1\, a^*_{Ly} = P^*_y \tag{3}$$

$$P_l\, a^*_{lx} + w^*_2\, a^*_{Lx} = P^*_x \tag{4}$$

$$r\, a^*_{kl} + w^*_2\, a^*_{Ll} = P_l \tag{5}$$

In addition to these, of course, we have altogether eight equations more involving eight flexible input coefficients except a_{Ly}, a^*_{kl} and a^*_{Ll}. Full employment conditions, on the other hand, imply

$$L_s = a_{sy}\, Y \tag{6}$$

$$l = a_{lx}\, X + a^*_{lx}\, X^* \tag{7}$$

$$L^*_1 = a^*_{Ly}\, Y^* \tag{8}$$

$$L^*_2 = a^*_{Lx}\, X^* + a^*_{Ll}\, l \tag{9}$$

$$K^* = a_{ky}\, Y + a^*_{ky}\, Y^* + a^*_{kl}\, l \tag{10}$$

$$L_e = a_{Lx}\, X + a_{Ly}\, y \tag{11}$$

These eleven equations determine the eleven variables[3]—P_l, r, W_s, W^*_1, W^*_2 and six quantity variables. The dichotomy between price and system and physical system holds. Given constant returns to scale technology equations (1) — (5) can determine P_l, r, W_s, W^*_1 and W^*_2 for given values of input-output coefficients which we automatically know once we know the commodity prices. It should be noted that the production structure in north has an input tier – output tier divisioning resembling, modified though, the Sanyal-Jones (1982) structure.[4]

Let us now consider different policies and investigate their influence on southern unemployment.

---

[3] In fact considering equations for input coefficents we have altogether nineteen equations to solve same number of variable. We do not make them explicit as that will simply increase the list of equations.

[4] One can indeed couch the entire analysis in terms of more simplified structures following Marjit (1990a, 1990B) and Jones & Marjit (1990). Our assumed structure seems to be more realistic, at the cost of simplicity, however.

Section – 2 : Policy

In this model any southern policy that aims at reducing unemployment has to affect production structure in north through movement of capital. This in turn affects I and export sector of south. A little algebra will show that.

$$\hat{I} = -\frac{1}{\lambda_{KI}^{*}}\left[\lambda_{KY}\ \sigma_{Y}\left(\hat{W}_{s} - \hat{r}\right) + \lambda_{KI}^{*}\ \sigma_{Y}^{*}\left(\hat{W}_{1}^{*} - \hat{r}\right)\right] \tag{12}$$

where λ's denote commodity shares of factor, σ's denote substitution elasticities and "hat" over a variable denotes proportional change, i.e., $\hat{x} \equiv \frac{dx}{x}$.

Equation (12) then represents the effect of a policy on northern production of I, thus affecting production of southern good –x. However, change in total employment can be computed only when we know exactly what happens to sectors producing x and y. In fact from (11) we get, given constant a_{LY},

$$\hat{L}_{e} = \delta_{1}\hat{X} + \delta_{2}\hat{Y} + \delta_{1}\hat{a}_{Lx}$$

where, $\delta_{1} = a_{Lx}X / L_{e}, \delta_{2} = 1 - \delta_{1}$. A little manipulation of (13) shows that,

$$\hat{L}_{e} = \delta_{1}\left[\alpha I + \gamma \hat{a}_{LX}^{*} + \beta \hat{P}_{I}\right] - \delta_{1}\hat{a}_{sY} + \delta_{1}\sigma_{X}\theta_{Ix}\hat{P}_{I} \tag{14}$$

where,

$$\left(\lambda_{Lx}^{*} + \lambda_{Ix}^{*}\lambda_{LI}\right) / \ \lambda_{Ix}\lambda_{Lx}^{*} = \alpha, \ \ \gamma = \lambda_{Ix}^{*} / \lambda_{Ix}\theta_{Ix}^{*}, \beta = \theta_{Lx}\sigma_{X}.$$

Equations (12) and (14) then set our stage to compare effects of different types of trade policies.

2.1. A Tax on Foreign Capital

Let us first consider a tax on foreign capital. This will raise capital cost of producing Y in south though capital cost in north will not change. Thus, $\hat{Y} < 0$ and $\hat{a}_{KY} < 0$. As it happens, foreign capital flows back to north raising production of I-good there and hence raising production of x in south. Therefore, employment of unskilled labour in southern y-sector falls while that in x-sector rises and a sufficient condition for employment increase wil be,

$$\delta_1 \alpha \frac{\lambda_{KY}}{\lambda^*_{KY}} > \delta_2 \tag{15}$$

By definition, $\alpha > 1$ so that the crucial factors underlying employment increase are the employment intensities in southern sectors, δ_1 & δ_2 and capital intensities in Southern Y sector, λ_{KY} and in northern input-tier, λ^*_{KI}. Even if the employment intensities in southern sectors are same, $\delta_1 = \delta_2$ and $\lambda^*_{KI} > \lambda_{KY}$, condition (15) can hold provided capital intensity in northern I-sector is not sufficiently larger than that in southern y-sector.

2.2 : A Tariff on Southern Imports

As an ad-valorem tariff on southern imported final good is imposed, price of Y increases. This via magnification effect pushes up W_S so that southern producers will substitute skilled labour by foreign capital,[5] i.e., $\hat{a}_{ky} > 0$ and $\hat{a}_{sy} < 0$. Capital flows in south thereby changing the production structure in north. Consider equation (10), from which we get.

$$\hat{K}^* = \lambda_{KY} \hat{Y} + \lambda_{KY} \hat{a}_{KY} + \lambda^*_{KY} \hat{Y}^* + \lambda^*_{KI} \hat{I} + \lambda^*_{KY} \hat{a}^*_{KY}$$

But as southern tariff leaves unchanged the northern commodity prices, by assumption, so are the factor prices. This means all input coefficents in north will be unaffected so that, given $\hat{Y}^* = -\hat{a}^*_{LY}$, we get.

$$\hat{K}^* = \lambda_{KY} \hat{Y} + \lambda_{KY} \hat{a}_{KY} + \lambda^*_{KI} \hat{I} \tag{16}$$

But as $\hat{K}^* = 0$, $\hat{Y} > 0$ and $\hat{a}_{ky} > 0$, so I must fall implying a fall in southern good-x. Change in employment is thus again ambiguous. However, if condition (15) holds then, however paradoxical it may seem, tariff unambiguously reduces southern employment. We find ourselves in a policy confict.

---

[5] We have ruled out any substitution possibility between skilled and unskilled labour.

2.3 An Export Subsidy on x

Suppose the government in south subsidises production of x. This will raise the domestic price of exporters by the amount of subsidy and hence given, $\overline{W}$ raises P_I. Thus, imported input will be substituted by unskilled labour so that emoployment will rise.

Again as, P_I rises, in north, production of I will go up and inrrespective of capital intensity, not only production of Y* but also of x* will fall. Production of x* will fall due to increase in cost of intermediate good as well as via Rybczynski effect causing greater release of L_2^* than absorbed in I-sector. Thus, W_1^* falls and this speeds up expansion of I-sector lowering cost of production. But as we have assumed fixed input co-efficient in I-sector so this means even greater demand for capital than released by Y*-sector, thus pushing up r. Foreing capital flows out of south reducing production of Y and employment. Net change in welfare is once again ambiguous.

Section –3 : Conclustion

We have tried in the above analysis to trace out the impact of conventional trade policies on aggregate employment in a typical LDC. It is observed that in all the cases the impact on employment is ambiguous as the policies affect the two southern sectors asymmetrically. What is interesting enough to find is that, the same condition which underlies the employment improving effect of taxing foreign capital or subsidising export production, yields diametrically opposite result when rariffs on final goods are imposed. Policies are in conflict then, as far as employment effect is concerned.

One policy that we did not discuss is wage subsidy in southern import good sector. This may be an interesting extension of the model.

Appendix

Consider equation (10) which can be written in percentage change form as,

$$\hat{I} = -\frac{1}{\lambda^*_{KI}}\left[\lambda_{KY}\hat{Y} + \lambda^*_{KY}\hat{Y}^* + \lambda_{KY}\hat{a}_{KY} + \lambda^*_{KY}\hat{a}^*_{KY}\right] \quad (17)$$

Similarly from (6) and (8) in turn we get,

$$\hat{Y} = -\hat{a}_{SY} \tag{18}$$

$$\hat{Y}^{*} = -\hat{a}^{*}_{LY} \tag{19}$$

On the other hand from the definition of elasticity of factor substitution, σ_j, we have

$$\hat{a}_{SY} - \hat{a}_{KY} = -\sigma_Y\left(\hat{w}_s - \hat{r}\right) \tag{20}$$

$$\hat{a}_{LX} - \hat{a}_{lx} = -\sigma_x \hat{P}_l \tag{21}$$

$$\hat{a}^{*}_{LY} - \hat{a}^{*}_{KY} = -\sigma^{*}_{Y}\left(\hat{w}^{*}_{1} - \hat{r}\right) \tag{22}$$

$$\hat{a}^{*}_{Lx} - \hat{a}^{*}_{lx} = -\sigma^{*}_{X}\left(\hat{w}^{*}_{2} - \hat{r}\right) \tag{23}$$

Using envelope property on the other hand,

$$\theta_{KY}\, \hat{a}_{KY} + \theta_{sY}\, \hat{a}_{sY} = 0 \quad \text{i.e. } \hat{a}_{sY} = -\frac{\theta_{Ky}}{\theta_{sY}} \hat{a}_{KY} \tag{24}$$

$$\hat{a}_{lx} = -\frac{\theta_{Lx}}{\theta_{lx}} \hat{a}_{Lx} \tag{25}$$

$$\hat{a}^{*}_{LY} = -\frac{\theta^{*}_{KY}}{\theta^{*}_{LY}} \hat{a}^{*}_{KY} \tag{26}$$

$$\hat{a}^{*}_{lx} = \left(\theta^{*}_{Lx} / \theta^{*}_{lx}\right) \hat{a}^{*}_{Lx} \tag{27}$$

Substitution of (18), (19), (20) and (21) in (17) yields,

$$\hat{I} = -\frac{1}{\lambda^{*}_{KI}}\left[\lambda_{KY}\, \sigma_Y\left(\hat{W}_s - \hat{r}\right) + \lambda^{*}_{KY}\, \sigma^{*}_{Y}\left(\hat{W}^{*}_{1} - \hat{r}\right)\right] \tag{17'}$$

Again as any change in output of I must be absorbed in x-good sector in both south and north so from (7) we notice,

$$\hat{I} = \lambda_{lx}\hat{X} + \lambda^{*}_{lx}\, \hat{X}^{*} + \lambda_{lx}\, \hat{a}_{lx} + \lambda^{*}_{lx}\hat{a}^{*}_{Lx} \tag{28}$$

Using (20) - (27) we get,

$$\hat{a}^{*}_{lx} = -\theta_{Lx}\, \sigma_x\, \hat{P}_l \tag{29}$$

$$\hat{a}^*_{lx} = -\theta^*_{Lx}\,\sigma^*_X\left(\hat{W}^*_2 - \hat{P}_I\right) \tag{30}$$

Again from (9),

$$\hat{x}^* = -\frac{\lambda^*_{Ll}}{\lambda^*_{Lx}}\,\hat{l} - \hat{a}^*_{Lx} \tag{31}$$

Thus, substitution of (27), (29) and (31) in (28) yields,

$$\lambda_{lx}\hat{X} = \hat{l} - \lambda^*_{lx}\left(-\frac{\lambda^*_{Ll}}{\lambda^*_{Lx}}\hat{l} - \hat{a}^*_{Lx}\right) + \lambda_{lx}\theta_{Lx}\sigma_X\hat{P}_I + \lambda_{lx}\frac{\theta^*_{Lx}}{\theta^*_{lx}}\hat{a}^*_{Lx}$$

$$\hat{X} = \alpha\,\hat{l} + \beta\hat{P}_I + \gamma\hat{a}^*_{Lx} \tag{32}$$

where,

$$\alpha = \left(1 + \lambda^*_{lx}\lambda_{Ll}\right)/\lambda_{lx},\ \ \gamma = \lambda^*_{lx}/\theta^*_{lx}\,\lambda_{lx},\ \ \beta = \theta_{Lx}\,\sigma_X.$$

Finally from (11), using (18), (21) and (32), we get

$$\hat{L}_e = \delta_1\hat{X} + \delta_2\hat{Y} + \delta_1\hat{a}^*_{Lx}$$

$$= \delta_1\left(\alpha\,\hat{l} + \gamma\hat{a}^*_{Lx} + \beta\hat{P}_I\right) - \delta_2\hat{a}_{sY} + \delta_1\sigma_X\,\hat{P}_I \tag{33}$$

Given these equations of change we turn to alternative policies discussed in the text.

A.1. Tax on Foreign Capital

Tax on foreign capital raises cost of capital, i.e. $\hat{r} > 0$ so that from equation –(2) in text we get,

$$\theta_{KY}\hat{r} + \theta_{sY}\,\hat{W}_s = 0$$

$$\text{i.e., } \hat{W} = -\hat{r}\frac{\theta_{KY}}{\theta_{sY}} \tag{34}$$

Since in north $\hat{r} = 0$ and $\hat{W}^*_1 = 0$ so using (17') and (34)

$$\hat{l} = \frac{\lambda_{KY}}{\lambda_{K^*l}}\delta_Y\frac{\theta_{KY} + \theta_{SY}}{\theta_{SY}}\hat{r} > 0 \tag{35}$$

Again

$$\hat{y} = -\hat{a}_{sY} = \sigma_Y \left(\theta_{KY} / \theta_{sY} \right) \hat{r} < 0 \tag{18'}$$

On the other hand as P_l remains constant, so are a_{LX}, a_{lX} and a^*_{LX}. Combining all these we get from (33).

$$\hat{L}_e = \delta_1 \; \alpha \, \hat{l} - \delta_2 \; \hat{a}_{sY}$$

$$= \hat{r} \left(\theta_{KY} / \theta_{sY} \right) \; \sigma_Y \left(\delta_1 \; \alpha \, \lambda_{KY} / \lambda^*_{Kl} - \delta_2 \right) + \delta_1 \; \alpha \, \lambda_{KY} \sigma_Y \hat{r} / \gamma^*_{Kl} \tag{36}$$

A sufficient condition for $\hat{L}_e > 0$ is therefore implied by condition (15) in the text.

A.2. Tariff on Y

Follow exactly the same way as in A.1. However the beginning change in price will now favourably affect W_s i.e., $\hat{W}_s > 0$.

A.3. Subsidy on X

Let ds > 0 be the amount of per unit subsidy on X. Then

$$\hat{P}_l = ds / \theta_{lX} > 0 \tag{37}$$

Now, from (4),

$$\hat{P}^*_X = 0 = \theta^*_{lX} \; \hat{P}_l + \theta^*_{LX} \; \hat{W}^*_2$$

$$\text{i.e., } \hat{W}^*_2 = -\theta^*_{lX} \; \hat{P}_l / \theta^*_{LX} = -\theta^*_{lX} \; ds / \theta^*_{LX} \; \theta_{lX} \tag{38}$$

Similarly from (5),

$$\theta^*_{Kl} \; \hat{r} + \theta^*_{LX} \; \hat{W}^*_2 = \hat{P}_l$$

$$\text{i.e., } \hat{r} = \left(\hat{P}_l - \theta^*_{LX} \hat{W}^*_2 \right) / \theta^*_{Kl} > \hat{P}_l > 0 \tag{39}$$

and from (3),

$$\theta^*_{KY} \; \hat{r} + \theta^*_{LY} \; \hat{W}^*_1 = \hat{P}^*_Y = 0$$

$$\text{i.e., } \hat{W}^*_1 = -\hat{\theta}^*_{KY} \; \hat{r} / \theta^*_{LY} < 0 \tag{40}$$

Substitution of (34) and (40) in (17') yields,

$$\hat{I} = \frac{1}{\lambda^*_{KI}}\left(\lambda_{KY}\,\sigma_Y\,\frac{\theta_{KY}+\theta_{sY}}{\theta_{sy}}\lambda^*_{KY}\,\sigma^*_Y\right)\hat{r} > 0 \qquad (41)$$

using (23) and (27)

$$\hat{a}^*_{Lx} + \theta^*_{Lx}\,\hat{a}^*_{Lx} / \theta^*_{Ix} = -\sigma^*_X\left(\hat{W}^*_1 - \hat{r}\right) = \sigma^*_X\left(\theta^*_{KY}\,\hat{r} / \theta^*_{LY} + \hat{r}\right)$$

i.e., $$\hat{a}^*_{Lx} = \frac{\theta^*_{Ix}}{\theta^*_{LY}}\sigma^*_X\,\hat{r} \qquad (42)$$

Therefore substitution of (18'), (37), (41), (42) in (33) yields,

$$\hat{L}_e = \delta_1\Big[\alpha\,\sigma_Y\,\hat{r}\,\lambda_{KY}(\theta_{KY}+\theta_{sY}) / \lambda^*_{KI}\,\theta_{sY} + \left(\alpha\,\sigma^*_Y\,\hat{r}\lambda^*_{KY} / \lambda^*_{KI}\right)$$

$$+(\beta ds/\theta_{Ix}) + \left(\gamma.\sigma^*_X\hat{r}\theta^*_{IX} / \theta^*_{LY}\right)\Big] - (\delta_2\sigma_Y\hat{r}\theta_{KY} / \theta_{sY}) + (\delta_1\sigma_X\beta ds/\theta_{Ix})$$

i.e.,

$$\hat{L}_e = \left[\left(\delta_1\alpha\,\sigma_Y\lambda_{KY} / \lambda^*_{KI}\right) - \delta_2\sigma_Y\right]\theta_{KY}\hat{r} / \theta_{sY} + \delta_1\Big[\left(\alpha\,\sigma_Y\lambda_{KY} / \lambda^*_{KI}\right)$$

$$+\left(\alpha\sigma^*_Y\lambda^*_{KY} / \lambda^*_{KI}\right) + \gamma\sigma^*_X\,\theta^*_{IX} / \theta^*_{LY}\Big]\hat{r} + (1+\sigma_X)\delta_1\beta ds/\theta_{Ix} \qquad (43)$$

Hence, again a sufficient condition for $\hat{L}_e > 0$ is given by (15) in text.

REFERENCES

1. Batra, R. and S Lahiri (1987), Imported Technologies, Urban Unemployment and the North-South Dialogue, Journal of Development Economics, 25, pp. 21-23.

2. Chan, K.S. (1982), The Employment Effects of Tariff under a Free Exchange Rate Regime, Jouirnal of International Economics, 8, pp 415-423.

3. Corden, W.M and R Findlay (1975), Urban Employment and Intersectoral Capital Mobility and Development Policy, Economica, 42, pp.59-78.

4. Jones, R W, (1956), The Structure of Simple General Equilibrium Models, Journal of Political Economy, 75, pp. 557-572

5. Jones, R.W, (1971), A Three Factor Model in Trade, Theory and History in J. Bhagwati et al. (eds.) Trade, Balance of Payments and Growth, North-Holland, Amsterdam

6 Jones, R.W. and S Marjit (1990), International Trade and Endogeneous Production Structures, Mimeo, University of Rochester.

7. Marjit, S, (1987), Trade in Intermediates and the Colonial Pattern of Trade, Economica, 54, pp. 173-184.

8 Marjit, S. (1988), Tariff, Technical Progress and Trade with Vertical Specialization, Keio Economic Studies, 25, pp 51-57.

9. Marjit, S. (1990 a), A Simple Production Model in Trade and its applications, Economics Letters, 32, pp 257-260.

10 Marjit, S (1990 b), Terms of Trade, Welfare and Capital Mobility in B. Dutta et al. (eds.) Essays on Economic Theory and Policy in honour of D. Banerjee. Oxford University Press

11. Sanyal, K.K. and R.W. Jones (1982), The Theory of Trade in Middle Products, American Economic Review, No 1, pp 16-32

12. Sauernheimer, K. (1986). Tariffs, Imported Inputs and Employment, Economica, 53, pp. 395-399.

International Debt, Terms of Trade Volatility and North-South Macroeconomic Relations

S. Mansoob Murshed and Somnath Sen

Preface

One of the co-authors of this paper – Somnath Sen – met Professor Prabuddha Nath Roy while lecturing at the University of Calcutta over two decades ago. Prabuddha-da, as he was then universally known, brought a degree of charm and elegance that was relatively uncommon to us in the hurly burly of Calcutta life. However, his placid exterior was a cover for a very active,inquisitive and dynamic mind that explored the intricacies of mathematical international trade theory to an extent that I have yet to meet in anyone else. Prabuddha-da was also a charming expositor and many a happy hour have I spent listening to his presentation. Many things have changed in my life since then, but I hope that some of the traits learnt then about teaching and research have lingered on. I am happy that this paper is written with my student – Mansoob Murshed – and hopefully I have succeeded in passing on some knowledge in the same way that I received from Prabuddha-da. I am proud to contribute to this volume.

I. Introduction

The 'debt crisis' suddenly emerged in 1982 when Mexico was about to default on the repayments of its international debt and since then a plethora of explanations for it have appeared. Some of these focus on the domestic macroeconomies of the Southern debtor nations, especially on their domestic monetary policies, the way in which deficits are financed, the need to pay for higher energy prices, and the possibility of capital flight to Northern financial markets. Other approaches point to the excessive availability of credit in the Northern financial sectors in the mid-1970s and often negative real rates of interest, leading to excessive lending. Still others examine the on-going process of debt rescheduling from the point of view of commercial banks, attempts to avoid a liquidity crisis and debtor nations' wishes to reduce the burden of debt servicing.

An alternative approach is to explain the debt crisis as a further and new manifestation of the nature of North-South economic interaction: the structuralist explanation. As such this approach is embedded in the nature of exchange between fix-price (Northern) and flex-price (Southern) commodities. From the structuralist position one would view the crisis in Southern indebtedness, or its conversion into a debt trap, as greatly a consequence of contractionary northern policies. Monetary policy is one of the major policy instruments available to the North, others include protectionism. The causal link, however, between Northern contraction and Southern indebtedness, which also engenders a global recession, is falling commodity prices. It is this very collapse in the Southern terms of trade, coupled with the increased difficulty it encounters in exporting to the North, which *intensifies* the position of Southern indebtedness. The tight macroeconomic stance of the North and the sharp decline in commodity prices of the early 1980s cannot be viewed as merely occurring simultaneously, the former is a major cause of the latter, leading on to a debt crisis, as will be *emphasised* in our analysis.

The economic relationships between North and South have usually been described as possessing 'asymmetric interdependence' or 'non-reciprocal dependence'; which is a structuralist tradition, dating back to Prebisch (1950) and more recently discussed in Taylor (1983). In this setting one of the major asymmetries in North-South interaction is the manner in which commodity prices feed into the Northern inflationary process [see Kaldor (1976) and Kanbur and Vines (1986)]. Basically, the prices of Southern commodities are determined mainly by Northern demand. Given that this is so, Southern prices or terms of trade are allowed to decline, during downswings in Northern economic activity; but are prevented from rising during upswings or booms in the North. This asymmetry arises due to the phenomenon of real wage resistance in the North. If the Northern consumption basket includes Southern goods, excess demand for these causes rising Southern prices, but also leads to wage and price increases in the North, preventing a rise in the Southern terms of trade. More importantly, excess demand for Southern goods can lead to steady-state inflation proceeding simultaneously in both regions, as will be shown in the model which follows. This, obviously, can only occur during boom periods in the world economy, generated by the North. These booms are unsustainable due to the inflationary implications of Northern real wage resistance (the attempt to maintain constant real wage and profit rates in the North). In the ensuing deflationary process not only do Southern terms of trade decline,[1] but a Southern debt crisis can also emerge, as the North shifts part of the burden of disinflation to the South. These periods are associated with a slowdown in the global economy. We concentrate on the nature of the global business cycle, rather than examining long-term trends in the Southern terms of trade, say. Such cycles of booms, with rising commodity prices and Northern inflation followed by slumps, broadly fit the stylised facts of post-war movements in the Southern of trade vis a vis the North. In the post-war period, Southern nations have felt an increasing need to import from the North, often through persistent balance of payments deficits [see World Tables (1983)], implying the possibility of debt crises emerging.

It is useful at this stage to enumerate some of the major asymmetries in North-South interactions, as employed in our model. The first major asymmetry arises from the structuralist notion that in the short run at least the North exhibits Keynesian[2] features while the South is classical (Walrasian) in nature. This means that demand determines *output* in the North and *prices or the terms of trade* in the South. When we combine this feature with the fact that in the South consumption varies in fixed proportion to income and not relative prices, the terms of trade for the South can be entirely determined by Northern demand. The South, however, has no influence over Northern prices, which are determined by Northern cost considerations. Another major asymmetry in North-South relations is the fact that the North being Keynesian, benefits from an export multiplier effect in the short run, whereas the South does not benefit from any such effect. We make a fundamental distinction, following the classification proposed by Hicks, between fix-price and flex-price markets. In the context of North-South models, the basic structuralist assumption is that the North obeys a Keynesian quantity adjustment process whereby excess effective demand is met by increased output. The aggregate price level in the North responds to cost push factors, in particular the strength of workers resistance to real wage cuts. On the other hand, the South operates within an auction market system and increases the price of its product when there is excess demand.[3] The nature of the goods produced in each region, the form in which specialisation has taken place historically, the cost of holding inventories, the competitiveness or monopolistic nature of the markets, all dictate the choice of flex/fix price structures for the two agents.

The paper is organised as follows. The next two sections specify the model, in terms of the equilibrium relations and the medium–term dynamics. Section IV discusses the implications, for debt and commodity prices (Southern terms of trade), of contractionary monetary policy and interest rate rises in the North. Section V explores the role of anticipations, specifically the change in dynamic movements of the variables of interest, if the South is able to anticipate the shocks emanating from the North. Section VI discusses an area of related interest, the role of changes in labour productivity and supply shocks in a debt-ridden world economy. The final section summarises and concludes.

II. The Model

The North is Keynesian where output or aggregate demand, Q_I is composed of absorption, A; autonomous expenditure G; the trade balance, exports, $c_s[1- v_s(D)] P_A \overline{Q}_A$ minus imports, $P_A M$; P_I is the price of the Northern good, P_A is the price of the Southern good, and $\sigma = P_A/P_I$, the terms of trade. After normalising the Northern equation by P_I, we obtain:

$$Q_I = A(Q_I + D; \sigma; r) + G + \sigma c_s[1-v_s(D)] \overline{Q}_A - \sigma M(Q_I+D; \sigma) \quad (1)$$

where $A_1 > 0$, $M_1 > 0$, $M_2 < 0$, $A_1 > \sigma M_1$, $A_3 < 0$, r is the Northern or world rate of interest. Northern income rises with the holding of the debt instrument, D (which is Southern debt). We assume that the Marshall-Lerner condition holds:

$$c_s[1- v_s(D)] \overline{Q}_A - M - \sigma M_2 > 0$$

$A_2 > 0$, by the Laursen-Metzler (1950) effect which gives us the Laursen-Metzler condition that $A_2 = M(1- \epsilon)$, where ϵ is the elasticity of absorption with respect to real income. [A detailed derivation of this condition can be found in Dornbusch (1980)]. The intuition behind this notion is based upon a concept of real income and expenditure which is obtained by deflating nominal income by a composite price level, which is a weighted product of the domestic and foreign price levels. Thus, if the terms of trade improve, P_a falls, then there is an increase in real income and consequently absorption. However, since the marginal propensity to absorb (consume) is less than unity only a part of the increased income is absorbed. Nominal absorption (in terms of home prices) will decline. This will operate as an adverse aggregate demand effect. It is worthwhile observing that in most of the literature on North-South interaction, attention on the terms of trade is focussed almost exclusively on the South. We bring it more firmly into the centre of things by giving it due emphasis in the *Northern* model. This constitutes our major argument for employing the Laursen-Metzler effect. Furthermore, in a Keynesian economy, when expenditure is divided up between domestic and imported goods, the price of the former being fixed in the short run and the latter variable, we have a strong case for adding the terms of trade as an argument in the absorption function.

We assume that output or income is demand determined, such that it is always equal to aggregate demand in equilibrium. This allows us to rewrite (1) as:

$$Q_I = F\ (\sigma, D,\ r) \tag{2}$$

where $F_1 > 0$; $F_2 > 0$, and $F_3 < 0$, all from (3).

The equilibrium relation for the South is modelled, following Kanbur and Vines (1986) by postulating a linear expenditure system. Thus:

$$P_A\overline{Q}_A = c_s P_A\overline{Q}_A + P_A M\{F(\sigma, D, r) + D; \sigma\} - c_s\{1 - v_s(D)\}P_A\overline{Q}_A \tag{3}$$

The total propensity to consume, c_s, (out of national income, akin to GDP, say) is fixed – the rationale behind it are the low levels of income in the South, making consumption proportional to income, but invariant with relative price changes. The shares of the home good and the imported good are v_s and $1-v_s$ respectively; v_s is made to vary positively with D, the stock of debt, which is issued to finance the Southern current account deficit, as will be explained below. Southern output, $\overline{Q}_A$ is held exogenously fixed (otherwise it can be made a positive function of its own price or terms of trade); P_A is the price of the Southern good; M is the quantity of Southern exports (Northern imports); and $c_s\{1 - v_s(D)\overline{Q}_A\}$ are Southern imports (Northern exports). Normalising (3) by P_A and cancelling terms gives the equilibrium relation for the South:

$$\overline{Q}_A = c_s\, v_s\, (D)\, \overline{Q}_A + M\, [F\, (\sigma, D, r) + D;\, \sigma\,] \qquad (4)$$

Another major feature of the Northern economy is the phenomenon of real wage resistance. This is seen to imply that workers are concerned with a particular standard of living or *level* of real wages which is defined in terms of a consumption basket, containing both domestic and Southern goods. The cost of purchasing this basket will be given by a price index, P, where $P = P_I^{\ \delta}\, P_A^{\ 1-\delta}$, δ and $1-\delta$ are the expenditure shares of the Northern and Southern goods, respectively. Northern firms follow a mark-up pricing rule for the Northern price, P_I over wage costs W, given the output-labour ratio (a) and desired profit rate of, g, over costs:

$$P_I = W(1+g)/a \qquad (5)$$

We assume that a and g are fixed parameters. Workers in the North have a targeted real wage, $\overline{w}$, related to the cost of living index, P. Their money wage, W is thus:

$$W = \overline{w}P$$

Using the definition of P and substituting the value of W in (5), we obtain a *desired* level of the terms of trade, $\overline{\sigma}$. Note that the desired level of $\overline{\sigma}$ depends on $\overline{w}$ [see Dornbusch (1980) for details]. When the actual and desired *real* wages converge, i.e. $w = \overline{w}$, then clearly the actual and desired terms of trade are equalised. This implies no wage cost-push inflation in the North. Consequently, for price equilibrium in the North:

$$\sigma = \overline{\sigma} \qquad (6)$$

On the other hand, if the actual real wage is less than its desired level, $w < \overline{w}$, then attempts by workers to raise money wages would lead to rising Northern prices and positive inflation rates. This implies that when $\sigma > \overline{\sigma}$, P_I would be rising due to the mounting pressure of wages.

We assume the North runs a balance of trade surplus with the South. This fits the stylised facts [see World Tables (1983)]. The Northern trade surplus, leads to the accumulation of debt by the South, which it uses to finance its trade deficit. Following Rodriguez (1979) we introduce a (debt) service account into the current account. A debt instrument, D, akin to a bond (in perpetuity) pays one unit of P_I per period. The rate of interest, r, is the North's interest rate, assumed exogenous, since we focus attention on real, structural features. It will rise when tight monetary policy is pursued. The price of D = 1/r. Thus, the value of the stock of bonds is also the value of interest payments on it.

The rate of accumulation of the bond is given by the current account surplus of the North [from (1), (2) and (4)]:

$$\frac{D}{r} = \sigma\, c_s\, (1 - v_s(D)]\, \overline{Q}_A - \sigma\, M\, [F(\sigma, D, r) + D\, ;\, \sigma) + D \qquad (7)$$

where $\dot{D}/r$ = rate of change of the value of bonds and rD/r = Northern value of interest payments.

In the steady-state, $\dot{D}/r = 0$, i.e. the rate of change of the Northern (Southern) current account surplus (deficit) is zero (constant). From (7) a rise in the stock of debt, D, or a rise in interest payments on debt r, will improve the Northern service account [last term of (7)]; hence for $D/r = 0$, its trade account must deteriorate. But, from (1) and (2) it is clear that an improvement in the service account enhances Northern income as D rises, thus imports, M will rise. The South is unable to affect the Northern import function (its exports) – which is determined by Northern variables, one of the asymmetries in North-South relations. But the maintainence of equilibrium requires an unchanged value of the current account $\dot{D}/r = 0$). Thus, when Northern imports rise as D rises, its exports must fall to keep up a constant value of the current account surplus. Northern exports (Southern imports) are $c_s[1 - v_s(D)]\, \sigma\, \overline{Q}_A$. Given the short run fixity of Southern output, Q_A and the aggregate propensity to consume c_s, the share of imports in Southern consumption, $1 - v_s$ must fall as D rises. In other words $v_s' > 0$ if $(1 - v_s') < 0$, when D rises. This squares in with our intuition and the stylised facts. A rise in the volume of debt (and debt service) necessitates some form of belt-tightening. The obvious way to do this is to import less and alter behaviour patterns such that the import content of consumption falls.

III. Medium Term Dynamics

In the North quantities adjust such that output is always equal to aggregate demand. Furthermore, we assume that when the actual terms of trade, σ, exceeds the required terms of trade, $\bar{\sigma}$ in (6), consistent with workers' desired standard of living, Northern workers obtain higher money wages, which leads to higher P_I as costs to firms increase :

$$\frac{P_I}{P_I} = \frac{W}{W} = \beta\left[\sigma - \bar{\sigma}\right] \qquad (8)$$

(where a dot over a variable denotes its time derivative). In addition, we could have a constant 'core' rate of inflation added on to (8). This could be determined, for example, by the rate of monetary growth. For notational simplicity we assume that the 'core' rate is zero.

In the South we assume that excess demand or supply leads to price adjustments. Hence the Southern goods market can be said to have a Walrasian flex-price structure in contrast to the North which exhibits fix-price characteristics. Given the widespread prevalence of futures markets, and the fact that the South requires to sell its product relatively quickly due to high inventory costs, we assume that the *expected* (e) rate of Southern price inflation is a function of excess demand for its output. Thus :

$$\left[\frac{\dot{P}_A}{P_A}\right]^e = \alpha\ [c_s\ v_s\ (D)\ \bar{Q}_A - \bar{Q}_A + M\ (\cdot)]$$

Some assumption needs to be made about the expectations formation mechanism. The optimum is to claim that the South has rational expectations so that it uses all the available information to forecast price changes. Hence systematic errors are avoided –a reasonable statement given the growing sophistication of international commodities markets (and Southern suppliers). In the absence of uncertainty, we get perfect (myopic) foresight and hence actual inflation is :

$$\left[\frac{\dot{P}_A}{P_A}\right]^e = \alpha\ [c_s\ v_s\ (D)\ \overline{Q}_A - \overline{Q}_A + M\ \{F(\sigma,D,r) + D;\ \sigma\}] \quad (9)$$

What of the terms of trade? We have defined it as $\sigma = P_A/P_I$. Thus changes in the terms of trade would be given by $\dot{P}_A/P_A$ minus $\dot{P}_1/P_I$ or (9) – (8). Thus

$$\dot{\sigma}/\sigma = \alpha\ [c_s\ v_s\ (D)\ \overline{Q}_A - \overline{Q}_A + M\ \{F(\sigma,D,r) + D;\ \sigma\}] - \beta[\sigma - \overline{\sigma}] \quad (10)$$

Given that Q_I, σ are the variables of most interest, (10) implies the possibility of steady-state inflation. This could occur if there was excess demand for the Southern good in the North, say. This would mean rising Southern prices i.e. $\dot{P}_A/P_A > 0$ in (9). But rising costs of the Southern good, would lower the Northern workers standard of living, leading to higher money wage demands. Firms will pass these increased wage costs in the form of higher prices. Thus $\dot{P}_1/P_I > 0$ in (8). But an inspection of (9) indicates that $\dot{\sigma}/\sigma = 0$, i.e. the terms of trtade are constant if $\dot{P}_1/P_I = \dot{P}_A/P_A$.In other words, we could have a steady-state with $\dot{\sigma}/\sigma = 0$, but positive inflation rates in both regions as $\dot{P}_1/P_I$ and $\dot{P}_A/P_A$ were both positive. This is similar to the analysis of Cardoso (1981), examining the inflationary process in the context of a two-sector, agriculture-industry model of a single country.

Steady state 'equilibrium' is given by setting (10) and (7) to zero. Given the possibility of positive and equal inflation rates in both regions in the steady-state, it cannot be called a full long-run equilibrium, and as such is a short to medium-run phenomenon. At the steady-state there could be excess demand for the Southern commodity, causing the inflation, and any targeted real wage is continuously being frustrated, through the chain of money wage increases leading to Southern price, P_A increases, back to nominal wage and price increases in the North. This steady-state occurs when the Southern goods market is not clearing – hence the inflation in both regions.

We can now evaluate parameter changes, such as in r, $\overline{\sigma}$ and $\overline{Q}_A$. Totally differentiating (10) and (7) and arranging them in matrix form (setting dG = 0) :

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} d\,\sigma^* \\ d\,D^* \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \begin{bmatrix} d\,r \\ d\,\overline{\sigma} \\ d\,\overline{Q}_A \end{bmatrix} \qquad (11)$$

where :

$a_{11} = \alpha\ M_1\ F_1 + \alpha\ M_2 - \beta$

$a_{21} = c_s\,[1 - v_s\,(D)]\ \overline{Q}_A - \sigma\ M_2 - M - \sigma\ M_1\ F_1$

$a_{12} = \alpha\ [c_s\,v_s'\ \overline{Q}_A\ M_1\ (1 + F_2)]$

$a_{22} = 1 - c_s\,v_s'\ \sigma\ \overline{Q}_A - \sigma\ M_1\ (1 + F_2)'$

$b_{11} = -\ \alpha\ M_1\ F_3$

$b_{21} = \alpha\ M_1\ F_3$

$b_{12} = -\ \beta$

$b_{22} = 0$

$b_{13} = \alpha\ [1 - c_s\ v_s\,(D)]$

$b_{23} = -\ \sigma\ [c_s\,\{1 - v_s\,(D)\}]$

The trace of the Jacobian cannot be signed unambiguously; $\alpha > \beta$ is a necessary condition for the determinant to be negative. This points to the possibility of saddle-path instability.

The medium-run dynamics can be depicted diagrammatically in σ, D space :

$$\frac{d\,\sigma}{d\,D}\,/\dot{\sigma}\,/\,\sigma = 0 = \frac{a_{12}}{a_{11}} \qquad (12)$$

$a_{12} > 0$, the necessary condition for $a_{11} > 0$ is that $\alpha > \beta$, the speed of adjustment of Southern prices is more rapid than in the Northern case. (We expand on this below.) If $a_{11} > 0$, then $\sigma/\sigma = 0$ curve is negatively sloped, from (10), depicting constant terms of trade.

The relation indicating a constant current account surplus (for the North) from (7) is :

$$\frac{d\sigma}{dD}/D=0 = \frac{a_{22}}{a_{21}} \tag{13}$$

$a_{21} > 0$ if the Marshall-Lerner condition is stronger than the Laursen-Metzler effect. When the terms of trade deteriorate for the North, σ rises, its trade balance improves by the former effect; while imports rise (trade balance falls) because of the positive absorption effect a la Laursen-Metzler. We assume the former dominates the latter, $a_{21} > 0$; $a_{22} < 0$ if $d\dot{D}/dD < 0$ in equation (7), as in Rodriguez (1979). This appears valid for a creditor region, like the North, bearing in mind that the rate of change of D is formulated in terms of the current account surplus. The creditor region's wealth is augmented by an increase in D, it imports more (M rises), but exports $[c_s\{1 - v_s(D)\}\overline{Q}_A]$ fall, so in balance $d\dot{D}/dD$ and a_{22} are likely to be negative. Hence the $\dot{D} = 0$ curve is positively sloped. The $\dot{\sigma}/\sigma = 0$ and $\dot{D} = 0$ curves are depicted in figure I along with the saddle-path SS. This is the phase diagram of the system.

We postulate that the speed of adjustment of Southern prices, α, is greater than that for the North, β. One reason for this holding, is the presence of implicit and explicit wage contracts in the North. It is costly to renegotiate wage contracts continuously [see Fischer (1977) and Dornbusch (1980)]. These are likely to make wage responses in the North to changes in the cost of living basket P, occur with lags; and in turn the change in Northern price, P_I, which is a mark-up on wage costs, will be somewhat sluggish from (8).

On the other hand, it is our contention that the Southern good is marketed in a more competitive environment, which is partially reflective of the weaker market power of the South, and even when the commodity in question is a manufactured product.

Given the nature of the model, the equilibrium (point E in Figure 1) is a saddle-point. The unique path which leads the system on to E is given by the saddle-path SS. Stability analysis therefore becomes complicated and how the economy reaches SS, from any initial point, is important. Faced with this problem, rational expectations models assume that the endogenous variables of interest (here D, σ) can be divided into two catergories: backward-looking or predetermined and forward-looking or non-predetermined. The choice of variable classification is a matter of economic intuition. We assume that the stock of debt is given at any point of time by its historical evolution; thus it is predetermined. Let the initial debt be D_0 (see figure 1).Northern price P_I is settled by wage contracts and are effectively given, currently. Then stability can be assured if and only if P_A, the Southern price can jump. This will allow the terms of trade σ to change discontinuously (jump) to attain the stable path SS at the level σ_0 (point F in figure 1). P_A (and hence σ) is a non-predetermined variable, able to respond to current information and news. This forward looking behaviour of South's price is plausible, particularly when it is in Southern interests to desire equilibrium in the global macroeconomy, in addition to the competitive nature of the market for the Southern product. The alternative to equilibrium for the South can be catastrophic, with $D \rightarrow \infty$ or $\sigma \rightarrow 0$. The former implies an extremely adverse wealth effect, as the stock of Southern debt tends to infinity; while the latter implies the complete inability to import any Northern goods. In the context of North-South relations, characterised by asymmetries and non-reciprocal dependence, P_A is a likely candidate for a jump variable, which is forward looking.[4] This is yet another implication of the South's weaker power in an asymmetrical international trading system of unequal exchange.

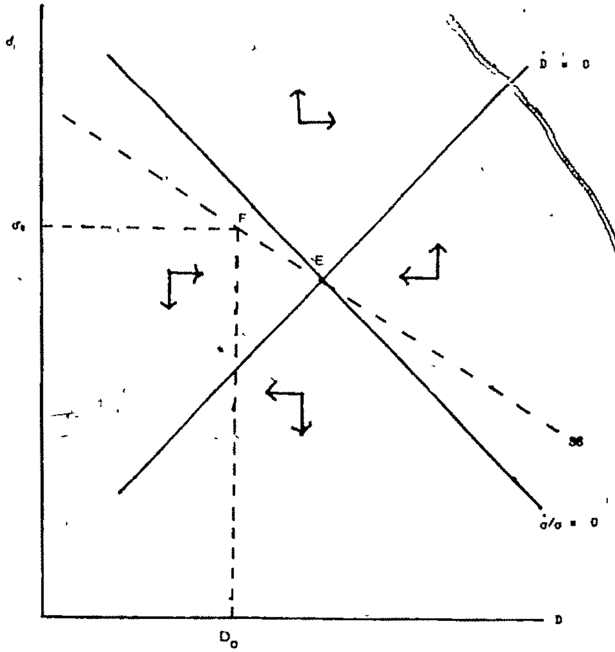


Fig.-1

IV. Contractionary monetary policy in the North.

The steady-state equilibrium described by equations (10) and (7) can be associated with positive inflation in both regions, emanating from excess demand for Southern commodities. For the North, inflation could be an undesirable phenomenon, except in the short run. One response, for the North is to pursue a contractionary monetary policy, to reduce or eliminate inflation, allowing for a steady-state with reduced or zero inflation (lower or zero excess demand for Southern commodities). Contractionary monetary policy in the context of our model implies an exogenous rise in Northern (world) interest rates. The changes in equilibrium σ^*, and D* are :

$$\frac{d\sigma^*}{dr} = \frac{-\alpha M_1 F_3}{|J|} < 0 \text{ as } |J| < 0 \tag{14}$$

and

change at the next point of time. There is a discontinuity in the dynamic movement of the terms of trade through time. When the (unanticipated) shock actually occurs and Southern producers realise that their demand will fall they reduce P_A rapidly (instantaneously) to attain the stable path SS_1. The dynamic movement of the system is characterised by a rapid fall in the terms of trade (through the jump in the Southern price) from E_1 to F; subsequently the system moves along the stable path from F towards E_2. In effect the Southern relative price falls quickly and then slowly declines towards its new steady-state equilibrium value. Debt rises continuously. The collapse of commodity prices and the increase in debt burden are the real life counterpart of the stylised process described here.

A rise in the Northern (world) interest rate, r, will raise the burden of debt and debt repayments, rD for the South. Therefore, it will be able to import less from the North. From (2), Northern absorption will decline as r rises, its exports to the South will fall also, all of which eases the excess demand for Southern commodities. The terms of trade, σ will decline rapidly. Not only is the Northern inflation rate reduced as excess demand for Southern goods falls, when Northern absorption declines; but more importantly, the pressure of real wage resistance is relaxed in (8) as $\sigma - \bar{\sigma}$, the gap between the actual and desired levels of the terms of trade declines. This is an important channel for lowering Northern inflation, as this was the very mechanism through which excess demand for Southern goods, led to Northern, P_I, price inflation. In a world economy, where the North has trading links with the South, it can *shift* part of the burden of disinflation to the South, via the fall in Southern terms of trade and the accretion of debt. In other words, the output (and employment) costs of disinflation to the North, would be higher, had it not been for the South bearing part of the burden of adjustment. We believe this process ties in with the stylised facts of the 1980s, when contractionary monetary policy was pursued in OECD countries to combat inflation, which led to a collapse of commodity prices and the emergence of the 'debt crisis'. It is our contention that the debt crisis is a *symptom*, of the partial shift of the burden of disinflation to the South [see also Murshed and Sen (1987)]. The empirical analysis in Gilbert (1987) shows substantial linkages between the decline in the Southern terms of trade vis a vis the North and the incresed burden of debt servicing. These empirical findings lend support to our theoretical analysis.

V. Anticipated Shocks

The previous section emphasised the importance of (rational) expectations in analysing the behaviour of the terms of trade and the stock of debt in a North-South model. In particular, we demonstrated the effects of an *unanticipated* shock to the system. Contractionary monetary policy, which is initially unexpected, raises interest rates and causes the Southern terms of trade to jump downwards. Commodity prices collapse and the debt crisis is exacerbated. This explains the behaviour of the world economy in the early eighties.

The question we ask now is whether the analysis alters singificantly when the exogenous change in the Northern interest rate is *anticipated* by Southern sellers. This could happen if inflation begins to rise in the North and the expected response is contraction of monetary growth. Alternatively, a change in political scenario, with the advent of relatively more conservative governments, could herald a forthcoming adjustment. In principle, agents do learn from experience and the structural shifts that produce these exogenous shocks can sometimes be predicted. The question is how they can affect the model?

If the shifts are permanent then clearly expected or unexpected changes will make no difference to the long-run (steady-state) equilibrium values of the variables. What an anticipated shock will do is to alter the dynamic behaviour of the economy on the transition path and the traverse will be different from that analysed earlier.

Assume that, akin to a policy 'announcement effect', it is anticipated that the interest rate will rise T periods hence. Even though current r is the same, Southern sellers will take this additional information into consideration, particularly since they have rational expectations. Knowing that a global recession is forthcoming they will wish to sell more currently since future sales are going to be depressed. Hence they will lower their prices leading to a rapid fall in commodity prices once again. The terms of trade jumps downwards as before while the stock of debt at the current period is unchanged.

However, note that since agents have T periods to adjust the response will be more muted. Indeed, since the equilibrium, currently, is still at the old level E_1, a jump similar to that in Figure 2 is definitely destabilising. If the terms of trade did jump to the level given by F then the laws of motion will take the system away forever from the postulated new equilibrium at E_2 (see also Figure 1). The only way in this scenario, for achieving stability, is for P_A (or σ) to jump again when the actual policy shock occurs (T periods later). But it must be noted that a jump, or discontinuity, per se is not desirable. Price adjustments, based on optimising agents' demand and supply curves [see equation (9)] reflect continuous Southern price changes over time. When policy shocks do appear then producers are forced to adjust discontinuously to attain steady-states since that alternative (to long-run equilibrium) creates major problems (as explained in the last section). But this is something they would rather prefer not to do. Hence, any dynamic path will be characterised by a minimum number of jumps.

What then will the behaviour of the economy be when the interest rate change is anticipated to occur after T periods? Figure 3 shows the transition path from E_1 to E_2. To avoid cluttering the diagram we only use the two saddle-paths. The terms of trade falls rapidly as before but not to the level F. Its decline is therefore lower and at G. Since the change is 'instantaneous', at G the laws of motion pertain to E_1 (the old equilibrium). The stock of debt therefore declines and falls further. T periods hence the equilibrium shifts to E_2. The relevant saddle-path is now SS_2. To avoid multiple jumps the system should be on SS_2 by the intervening T periods. Hence the traverse from G to H should take T periods also. The size of the fall E_1 G is determined by the length of T. If events can be anticipated very much earlier then the jump is less and the time taken to move along GH is correspondingly longer. On the other hand, if T is short then the new saddle-path needs to be attained quicker; hence a larger relative collapse of the terms of trade.

When the actual policy change does take place the economy is on SS_2 and hence moves along the saddle-path towards its new equlibrium. The actual movement is therefore E_1 G, GH, and finally H to E_2. The initial reduction in the terms of trade, without the corresponding recession in the North engendered by the rise in interest rates, causes some debt retirement from Southern export promotion. Unfortunately, when contractionary policy actually bites then the debt burden rises again over time. The only advantage that the South has is that there is more time for adjustments. Hence, the 'collapse' is less and may not be a disaster. A viable system of early warning and price monitoring would therefore be useful since it will allow the South to initiate domestic policies concurrently to counteract against these adverse shocks. In our global model, of course, we do not discuss such possibilities; but their importance remains. The effects of shocks are generally amplified when they are unanticipated.

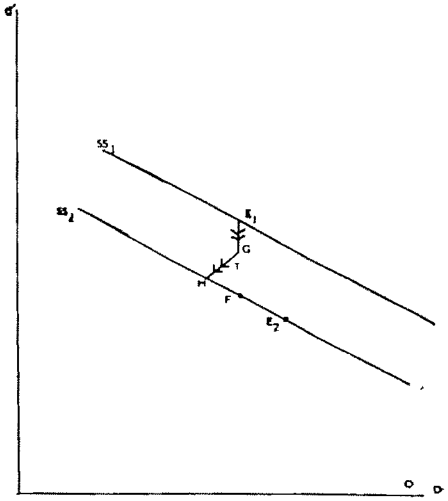


Fig.-3

VI. Productivity Shifts

A. *A Rise in Nôrthern Productivity*

From equation (5) above, an improvement in labour productivity, (a), lowers P_l for given W, as the cost of production declines. For any targeted level of real wages, $\overline{w}$, the implication is that the associated (required) terms of trade $\overline{\sigma}$ will rise, in (6) and (8), as the cost of living declines.

$$\frac{d\sigma^*}{d\bar{\sigma}} = \frac{-\beta\left[1 + c_s v_s{}' \ \sigma \bar{Q}_A + \sigma M_1 (1 + F_2)\right]}{|J|} \tag{16}$$

which is negative if $d\dot{D}/dD < 0$.

$$\frac{dD^*}{d\bar{\sigma}} = \frac{\beta\left[c_s \{1 - v_s(D)\} \ \bar{Q}_A - \sigma M_2 - M - \sigma M_1 F_1\right]}{|J|} \tag{17}$$

This is also negative, as long as the Marshall-Lerner trade balance effect dominates the Laursen-Metzler effect on absorption to the trade balance.

As the Northern price, P_I falls due to the productivity increase in (5), on impact it raises the terms of trade σ. The rise in the terms of trade, causes a decline in the Northern demand for Southern goods, hence the Southern price, P_A declines – in figure 4, P_A jumps to F from E_1, hence σ falls. At this point, due to the decline in the terms of trade, Northern demand for Southern commodities picks up and σ rises. But, the new steady-state is at E_2, with both a lower stock of Southern debt, D, and terms of trade, σ. The reason for this, is that the $\dot{D} = 0$ schedule from (7) is unchanged by the productivity increase. Thus, the rate of change in the Northern (Southern) current account surplus (deficit) is unchanged, although the *level* of D has fallen, as the Northern trade surplus (Southern deficit) has declined. The rise in Northern importation from the South, is constrained by the unchanged $\dot{D} = 0$ (7).

Another way of interpreting this result arises from the fact that the decline in the Southern terms of trade, has such a strong negative Laursen-Metzler effect on Northern absorption, as well as a substantial negative trade balance effect (via the Marshall-Lerner condition), that Northern aggregate demand and output declines. Along with it excess demand for Southern goods, thus the new steady-state value of σ is lower. Thus, *contrary* to the findings of Prebisch (1950), a Northern productivity increase will not benefit the South via improved terms of trade, although Southern indebtedness declines. Of course, the steady-state inflation rates in both regions will have declined.

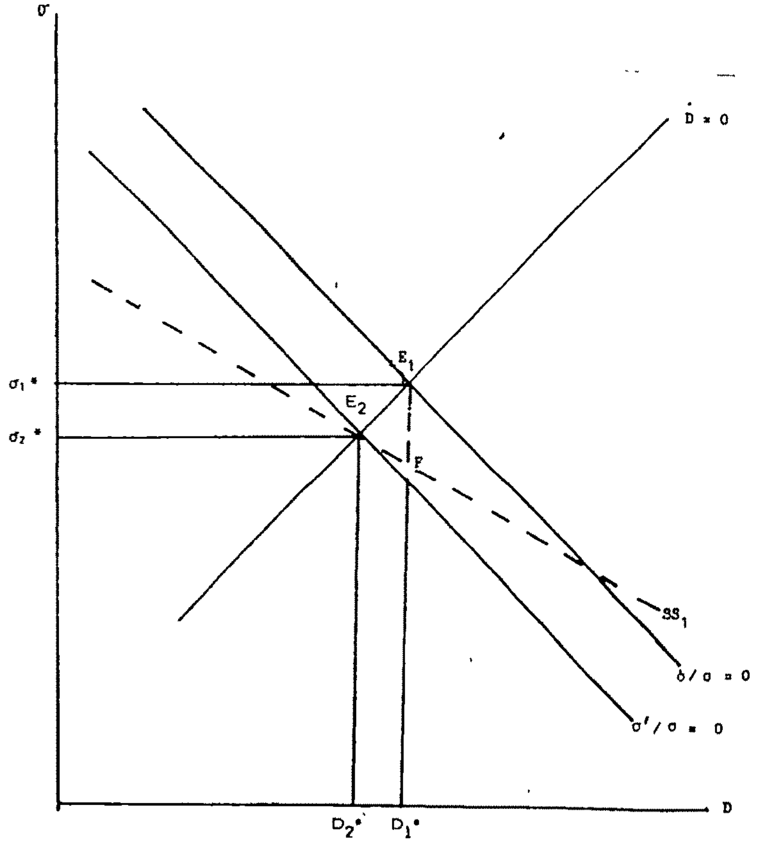


Fig.- 4

B. *A Rise in Southern Productivity*

In the context of our model, a rise in Southern productivity or technical progress in that region, is represented by an exogenous increase in Southern output, Q_A.

$$\frac{d\sigma^*}{d\bar{\sigma}_A} = \frac{\alpha[c_S - 1]\left[c_S v_S{}'\sigma\bar{Q}_A + \sigma M_1(1+F_2)\right] + \alpha\left[1 - c_S v_S\ (D)\right]}{|J|} \qquad (18)$$

This will be negative, if $c_s > 1$, i.e. as long as the South has a propensity to run a deficit (high propensity to absorb) divided

$$\frac{dD^*}{d\bar{\sigma}_A} = \left[1 - c_s\right] \left[\alpha\sigma(M_1 F_1 + M_2)\right] + \beta c_s \left[1 - v_s\,(D)\right] \sigma$$
$$- \alpha\left[1 - c_s v_s\,(D)\right] \left[c_s\left\{1 - v_s(D)\right\} \bar{Q}_A - M\right] \qquad (19)$$

by $|J| < 0$. Given our previous assumptions about $\alpha > 0$ and the existence of a Northern surplus, $dD^*/d\bar{Q}_A$ is likely to be positive.

The final outcome in this case, closely resembles the case of contractionary monetary policy, and figure 2 with the new steady-state at E_2 with lower σ and higher D applies. An increase in Southern output, is likely to drive down the Southern terms of trade, σ, as there is excess supply for the Southern good. At the new equilibrium the stock of Southern debt, D, will rise, if the Southern current account deficit (Northern surplus) increases. This is not unlikely at a higher level of Southern output, given its high propensity to consume i.e. when $c_s > 1$. As pointed out by Prebisch (1950), technical progress in the South, can make that region worse off via a decline in its terms of trade, which is compounded in this case by an increase in indebtedness.

Conclusions

To summarise, the debt crisis emerged as a consequence of Northern attempts at inflation control via tight monetary and other contractionary policies – raising world interest rates and the burden of debt repayment, accompanied by a sharp decline in the Southern terms of trade. The debt crisis is thus an additional and somewhat novel feature of the see-sawing relations between the (fix-price) North and the (flex-price) South as pointed out by authors such as Kaldor (1976), in the setting of real wage resistance in the North. The traditional structuralist result that productivity improvements in the South (Prebisch, 1950) harm that region via declining terms of trade is supported by our model. However, Prebisch's equally famous result that Northern productivity improvements benefit the South, is not totally supported by our model.

The short to medium-run nature of our model should be emphasised. We focus on the global macroeconomic business cycle. A 'full' long-run equlibrium would completely eliminate excess demand for Southern goods. In other words there would be a 'proper' mechanism for doing so. The medium term response to (steady-state) inflation, caused by excess demand for Southern goods, would be the use of contractionary policy instruments in the North. However, this only works in one direction, as it is implausible to expect expansionary policies to be pursued in the North during periods when Southern commodities are in excess supply. One manner in which the Southern market, could be made to clear in the long-run is by allowing quantities of goods supplied in the South to adjust. The South passively adjusts by varying quantities of say, some residual expenditure or basic needs consumption, while Northern factors determine both Northern output and the terms of trade. In the event of this occurring, the South becomes effectively *decoupled* from the world system, when it comes to terms of trade determination [see Murshed and Sen (1987)]. This decoupling of the South may be the feature of the *overall* post-war period taken as a whole.

The most immediate policy conclusion of our model emerges from the root of the problem - cyclical variations in the terms of trade, with its inflationary and counter–inflationary consequences. If Southern supply is related to its terms of trade, then commodity price stabilisation would smoothen out Southern supply and reduce excess demand for it. This constitutes a *macro* as opposed to microeconomic argument for commodity price stabilisation, as pointed out in Kanbur and Vines (1986). It would reduce the amplitude of the global business cycle; hence the output costs of disinflation internationally. However, when contractionary anti-inflation policies are (or have to be) pursued in the North, the costs of these policies for the South are mitigated somewhat if the South can anticipate them, as then it has more time to adjust. Thus, there are gains from an early warning of such events.

FOOTNOTES

1. Recent empirical evidence as in Beckerman and Jenkinson (1986), suggests that much of the success of recent anti-inflationary policies was due to the fact that the reductions in Northern demand, lowered commodity prices, but not because of moderation in the real wage resistance front. The model of Murshed and Sen (1987) demonstrates that the output cost of disinflation to the North is lower as it can pass on some of the cost to the South, via declining terms of trade.

2. The North may not be fully Keynesian in the context of intra-Northern trade, but we believe that a Keynesian paradigm is appropriate for the North, in analysing North-South, economic interaction.

3. To the extent that Southern nations, such as the NICs supply manufactured goods, even these are marketed in a broadly flex-price environment, due to the costs of holding inventories. Even if the production of manufactures exhibit certain fix-price characteristics in a *domestic* context, this will not be the case in trading links with the North A decline in demand from the North, will force Southern suppliers to lower costs via reductions in the real wage and / or real profit rates

4 A forward looking variable, such as the nominal exchange rate in the Dornbusch (1980) type models, is one whose laws of motion are influenced by current events, thus it will respond quickly to "news" and jump towards its 'desired' level The motion of a backward looking variable, on the contrary, is already determined by past events. The distinction is akin to that of control and state variables in optimal control models

REFERENCES

Beckerman, W.J. and Jenkinson, T. (1986). "What Stopped the Inflation? Unemployment or Commodity Prices", Economic Journal, 96, pp 39-54

Dornbusch, R. (1980) Open Economy Macroeconomics, Basic Books, New York.

Fischer, S. (1977). 'Long-term contracts, Rational Expectations, and the Optimal Money Supply Rule', Journal of Political Economy, 85, pp. 191-205.

Kaldor, N. (1976) 'Inflation and Recession in the World Economy', Economic Journal ', 86, pp. 703-14

Kanbur, S.M.R. and Vines, D. (1986). 'North-South Interaction and Commodity Control'. Journal of Development Economics, 23, pp. 371-87.

Murshed, S.M. and Sen, S. (1988) 'Northern Monetary Policy and the Southern Debt Crisis', paper presented at the European meeting of the Econometric Society, Bologna, August

Murshed, S.M. and Sen, S. (1989). 'Inflation and Macroeconomic Adjustments in a North-South Model', Greek Economic Review, 11, pp. 95-118

Prebisch, R. (1950) The Economic Development of Latin America and its Principal Problems, U.N , New York.

Rodriguez, C.A. (1979). 'Short-run and Long-run Effects of Monetary and Fiscal Policies under Flexible Exchange Rates and Perfect Capital Mobility', American Economic Review, 69, 176-82.

Taylor, L. (1983). Structuralist Macroeconomics, Basic Books, New York.

World Bank, World Tables, World Debt Tables and World Development Reports, various issues.

Dumping and Trade Policy under The Condition of Demand Uncertainty and Technological Inflexibility

Sarmila Banerjee

Section I : Introduction

The world recession since the mid 70s has brought the issue of dumping back to its prominence. Traditionally, dumping was considered as a practice of international price discrimination in segmented markets. However, the recession has created excess capacity in some industries and dumping has started taking place due to cyclical demand uncertainties and the imperfectly adjusting supplies. Thus, the relatively modern definition of dumping focuses on the practice of selling below cost in the foreign market. In this changed perspective it seems relevant to reassess the welfare implications of dumping for the importing country.

Banerjee (1989) has shown that when demands in both the countries are subject to cyclical fluctuations and the importing country faces factor-specficity together with factor-price inflexibilities, then under certain plausible conditions, dumping can be welfare reducing for the importing country. To capture the flavor of both traditional and modern definitions of dumping , we consider a situation where below-cost dumping is motivated by the profit possiblities of price discrimination, in a world with uncertain demand. This paper shows that for below-cost dumping in a world under uncertain demand, technological inflexibilities render the economic merit of dumping dubious. When dumping leads to the loss of economic welfare, that creates a scope for government intervention. The choice of efficient policy, obviously, depends on the nature of distortions leading to these inefficiencies.

The remainder of this paper is organized as follows. Section II briefly reviews the literature, section III describes the general features of the model, section III. 1 studies the nature of equilibrium in the exporting country and determines the optimal magnitude of below-cost dumping, when both home and foreign demands are stochastic. Section III. 2 analyzes the nature of equilibrium in the absence of dumping and then discusses the short-run and long-run effects of dumping on the importing country's equlirium price, output, and capacity decisions. Section IV examines the welfare effect of dumping on the importing country, whereas section V presents an efficiency ranking of alternative industrial policies. Finally, the concluding section presents the summary and conclusions.

Section II : A Brief Review of the Literature

Bernhardt (1984) has shown that domestic demand uncertainty in conjunction with adjustment costs in production might lead to below-cost dumping in a foreign market which has no uncertainty at all. In the face of home demand uncertainty, a firm has to choose its capacity level ex ante and any ex post deviation from capacity in either direction is costly. The home market is supplied by a discriminating monopolist who is a price taker in the competitive foreign market. The possibility of dumping raises both expected profit and expected consumer surplus in the exporting country.

Hillman and Katz (1986) studied the relationship between the exporting country's home market uncertainty, its form and nature, and the tendency to dump, i.e., a pro-dumping bias. The product in question is assumed to be perishable, so that inventory cannot be held to smooth out demand/supply fluctuations. It has been found that when home demand is uncertain the pro-dumping bias increases irrespective of the form of this randomness. On the other hand when home output (supply) is uncertain, with multiplicative uncertainty, the effect on the magnitude of dumping is ambiguous.

Das (1987) models both home and foreign demand which are subject to stochastic fluctuations in a dynamic context. Adjustment costs are infinite in the sense that within a particular time interval production decisions, once taken, cannot be revised. However, some flexibility has been allowed through the possibility of inventory management. The paper derives conditions for price-dumping and below-cost-dumping and the effect of increased correlation between home and foreign uncertainty parameters on the magnitude of dumping is studied.

All the above mentioned papers examined the issue of dumping from the exporting country's stand point. Our focus of analysis will be on the importing country, when the importing country also faces cyclical demand fluctuations and adjustment costs in production.

Section III : Essentials of the Model

There are two countries. Country I is the exporting (foreign) country and country II is the importing (home) country. Both home and foreign countries face cyclical demand uncertainties, produce a single homogeneous product and incur adjustment costs whenever ex post output decisions deviate from ex ante plans.

Assumption 1 : A single homogeneous product is produced in both countries and quadratic adjustment cost is incurred to revise output decision in the short-run. The general characterization of the cost function is as follows :

$$C(Q,K)_1 = \tau_1 Q + (\tau_2/2)\,(K-Q)^2 + F_1 \qquad ....(1)$$

$$C(Q,K)_2 = c_1 Q + (c_2/2)\,(K-Q)^2 + F_2 \qquad ....(2)$$

where $\tau_1, \tau_2, c_1, c_2 > 0$. Equations (1) and (2) represent cost structure faced by the firms in Country I and II respectively. Prior to the actual demand realization K is chosen. After knowing the demand firm can produce any amount. However, in addition to the constant unit variable cost $\tau_1(c_1)$ it incurs a quadratic adjustment cost to produce at a level other than K. The size of the short-run unit variable cost $\tau_1(c_1)$ can be thought of as a measure of the level of technology the firm possesses and $\tau_2(c_2)$ as a measure of flexibility of the technology. The quadratic, and therefore symmetric , nature of the short-run adjustment process means it is costly to contract production as well as to expand it. F_i denotes the level of fixed cost in country i,i = I,II. From equation (1) we have,$MC = \tau_1 - \tau_2\,(K-Q)$.Thus,$MC > \tau_1 (<\tau_1, = \tau_1)$ if $Q > K$ ($<K, = K$), respectively. The ex post average variable cost and average total cost for a given level of K are illustrated in figure 1.

Assumption 2 : In Country I there is a monopolist while in Country II the market is perfectly competitive. Demand is stochastic in both markets and the source of uncertainty is the business cycle.

Assumption 3 : There are only two possible states of nature, A and B, where A is the bad and B is the good state of demand. Both the states are equally likely, i.e. pr(A) = pr (B) = ½.

Assumption 4 : The random variable enters the linear demand functions additively.

$$P(Q) = -(\alpha / 2)\, Q + \beta + f \quad \text{where } f = \pm\Theta,\ E(f) = 0 \quad \ldots\ldots(3)$$

$$P(Q) = -(a / 2)\, Q + b + v \quad \text{where } v = \pm\delta,\ E(v) = 0 \quad \ldots\ldots(4)$$

where $\alpha > 0$, $\beta > \Theta > 0$, and $a > 0$, $b > \delta > 0$.

Assumption 5 : Both countries are in the same phase of business cycle together. This implies a positive correlation between f and v.

Due to the competitive structure in country II, the demand stochasticity will be perceived as a price stochasticity at the firm level. In the home market the monopolist is facing a demand shock while in the foreign market he is facing a price shock induced by a different demand shock. Unless supply is perfectly inelastic, for the same magnitude of a demand shock, the latter will change by less than the change in demand intercept itself. Thus the degree of uncertainty perceived at the firm level will be less in a competitive market than in a monopolistic one. In other words, in the foreign market, Country I is a price taker and, hence, in that market will not have to bear the entire demand shock alone; it will be jointly borne by all the firms in the industry. On the other hand in it's home market, being the sole producer, it absorbs the entire shock. Due to this asymmetry in market structure, we make the following assumption:

Assumption 6 : The home demand parameter (Θ) is more sensitive than the foreign price parameters (P_A, P_B) over the phases of business cycles.

Assumption 7 : Upon the realization of the actual state of nature country I exports to Country II's market whenever it is profitable. On the other hand Country II does not have the option of supplying Country I's market.

Suppose Q is the total production in Country I, then $Q = (Q_h + Q_e)$, where Q_h is the home sale, Q_e is the export. The monopolist will sell in both the markets when MR (Q) $< \bar{P}$, where

$\bar{P}$ is the expected foreign price and MR stands for marginal revenue. Since the foreign market is perfectly competitve, price equals marginal revenue there. The profit maximizing condition for a discriminating monopolist implies MC = MR $(Q_h) = \bar{P}$ = MR (Q_e). Now, at the capacity level of output, MC = τ_1 = AC. To concentrate on below-cost dumping we assume that $\tau_1 > P_B$. This implies that even in the event of the best possible demand realization, price in the importing country does not cover the average cost of production of the exporting country. Of course, it does cover the average production cost of the importing country. This essentially means:

Assumption 8 : Country II has some long run cost advantage over country I.

Since $\bar{P} < \tau_1$, in the absence of uncertainty the monopolist will have no incentive to export. But in a world with stochastic demand, with the above mentioned specification of the cost structure, marginal cost can be less than τ_1 for Q<K. Here, if the foreign price can cover marginal cost, then below-cost dumping might appear to be a profit maximizing outcome.

Assumption 9 : $(\beta - \Theta) > \tau_1 > P_B$.

The first part of the inequality implies the existence of $Q_h > 0$ at all possible states of nature whereas the last part implies, whenever $Q_e > 0$ it is below-cost dumping. Hence, dumping can be expected only in state A.

The determination of the optimal magnitude of dumping is discussed below. In this part of the analysis, we build upon bernhardt's (1984) work. We differ from Bernhardt in two respects. First, in addition to home demand uncertainty, foreign market is also assumed to be stochastic. Secondly, we have assumed that the random variable is binary rather than having a continuous density over a closed interval.

III. 1 : Exporting Country—the Profit maximizing Monopolist

In the absence of the possibility of foreign trade the risk neutral monopolist in Country I will maximize his expected profit by choosing a capacity level to satisfy expected demand. Hence, $k^* = E(Q_h)$. When the possibility of trade opens up the capacity can be expanded, since, in the event of good demand realization, that allows higher profits and in the case of bad demand realization, the marginal revenue will never go down below the foreign price. Thus, the possibility of below cost dumping raises expected monopoly profits provided that the foreign price is not too low in the bad state A. In state A, both f and v are negative. The negative f makes dumping attractive while the negative v makes it unattractive. If the positive influence of f is stronger than the negative influence of v then dumping will take place. Thus, $K^*_{autarky}$ can be expected to be smaller than K^*_{trade} and for a given capacity level $E(Q_{autarky})$ will be less than $E(Q_{trade})$. This point can be illustrated further in terms of figure 2.

In a bad state of demand, instead of choosing the output level $Q_{autarky}$ the monopolist can choose Q_T (for a given capacity level K^*) and can sell Q_h in home market where $Q_h<Q_{autarky}$ implying $MR(Q_h)>MR(Q_{autarky})$ and the remainder $(Q_T-Q_h)=Q_e$ can be sold in the foreign market at a price P_A. This price will not cover τ_1, the marginal cost at capacity level but will cover the marginal cost at the actual level of output. In fact, the profit maximizing condition is :

$$MC(Q)=MR(Q_h)=MR(Q_e)=P_A.$$

This intuition can be captured more rigorously in terms of the model specified above.

Autarky

The monopolist is choosing his optimal capacity by solving: $\max[½(\pi_A+\pi_B)]$ where π_i is the maximum profit in state i, i=A,B. for a given K. His cost and demand functions are given by equations (1) and (3) respectively.

$$\pi = -(\alpha/2)Q_h^2+(\beta+f-\tau_1)Q_h-(\tau_2/2)(K-Q_h)^2-F_1.$$

The first order condition gives,

$$Q_h^* = (\beta + f - \tau_1 + \tau_2 K)/(\alpha+\tau_2) \qquad ....(5)$$

Therefore, $Q^*_{h,A} = (\beta - \Theta - \tau_1 + \tau_2 K)/(\alpha+\tau_2)$

$$Q^*_{h,B} = (\beta + \Theta - \tau_1 + \tau_2 K)/(\alpha+\tau_2)$$

and $\pi_A = \{(\beta - \Theta - \tau_1 + \tau_2 K)^2 / 2(\alpha + \tau_2)\} - (\tau_2 / 2) K^2 - F_1$

$\pi_B = \{(\beta + \Theta - \tau_1 + \tau_2 K)^2 / 2(\alpha + \tau_2)\} - (\tau_2 / 2) K^2 - F_1$

Hence, $E(\max \pi) = ½ (\pi_A + \pi_B)$ and maximization over K gives the first order condition from which the optimal capacity can be derived as :

$K^*_{autarky} = (\beta - \tau_1)/\alpha$ ......(6)

It is probably not surprising to observe that

$E(Q^*_{autarky}) = ½(Q_{h,A} + Q_{h,B}) = (\beta - \tau_1)/\alpha = K^*_{autarky}$ ......(7)

and, $E(P^*_{autarky}) = ½(P_A + P_B) = ½(\beta + \tau_1)$ ......(8)

Trade

Let the monopolist have access to the foreign market. Then his optimization problem becomes,

$\max_{Q_h, Q_e} \pi_A = -(\alpha/2) Q_h^2 + (\beta - \Theta - \tau_1) Q_h + (p_A - \tau_1) Q_e$

$-(\tau_2 / 2)(K - Q_h - Q_e)^2 - F_1$ ......(9)

and $\max_{Q_H} \pi_B = -(\alpha/2) Q_H^2 + (\beta + \Theta - \tau_1) Q_H$

$-(\tau_2 / 2)(K - Q_H)^2 - F_1$ ......(10)

Recalling that the monopolist will have an incentive to sell in the foreign country only in state A the first order conditions of (9) are,

$$\begin{bmatrix} (\alpha + \tau_2) & \tau_2 \\ \tau_2 & \tau_2 \end{bmatrix} \cdot \begin{bmatrix} Q_h \\ Q_e \end{bmatrix} = \begin{bmatrix} \beta - \Theta - \tau_1 + \tau_2 K \\ p_A - \tau_1 + \tau_2 K \end{bmatrix}$$

from which we obtain the optimal state-contingent supplies,

$Q_h^* = (\beta - \Theta - p_A)/\alpha$ ......(11)

$Q_e^* = p_A(\alpha + \tau_2)/\alpha\tau_2 - \tau_1/\tau_2 + K - (\beta - \Theta)/\alpha$ ......(12)

The first order condition of (10) gives,

$Q_H^* = (\beta + \Theta - \tau_1 + \tau_2 K)/(\alpha + \tau_2)$ ......(13)

Substituting (11) and (12) into (9), we derive the maximum profit in state A as,

$$\pi_A^* = -(\alpha/2)[(\beta+\Theta-p_A)/\alpha]^2 + (\beta-\Theta-\tau_1)(\beta-\Theta-p_A)/\alpha +$$
$$(p_A-\tau_1)\,[p_A(\alpha+\tau_2)/\alpha\tau_2 - \tau_1/\tau_2 + K - (\beta-\Theta)/\alpha] - (\tau_2/2)$$
$$[-(\beta-\Theta-p_A)/\alpha - p_A(\alpha+\tau_2)/\alpha\tau_2 + \tau_1/\tau_2 + (\beta-\Theta)/\alpha]^2 - F_1$$

Similarly, substituting (13) into (10) yields,

$$\pi_B^* = -[\alpha/2(\alpha+\tau_2)^2].(\beta+\Theta-\tau_1+\tau_2K)^2 + (\beta+\Theta-\tau_1)$$
$$(\beta+\Theta-\tau_1+\tau_2K)/(\alpha+\tau_2) - (\tau_2/2)$$
$$[K-(\beta+\Theta-\tau_1+\tau_2K)/(\alpha+\tau_2)]^2 - F_1$$

From these profit functions, the optimal capacity follows.

$$K^*_{trade} = (p_A-\tau_1)\,(\alpha+\tau_2)/\alpha\tau_2 + (\beta+\Theta-\tau_1)/\alpha \qquad \text{.......(14)}$$

Substituting (14) for K in (12) and (13), we get,

$$Q_e^* = 2(p_A-\tau_1)\,(\alpha+\tau_2)/\alpha\tau_2 + 2\Theta/\alpha \qquad \text{.......(15)}$$

$$Q_H^* = (\beta+\Theta+p_A-2\tau_2)/\alpha \qquad \text{.......(16)}$$

where, Q_H^*: total output in state B

$Q^*_{trade,A} = (Q_h^*+Q_e^*)$: total output in State A.

Comment 1:

The necessary and sufficient condition for dumping in state A is, MR $(Q^*_{trade,A}) < p_A$ or $\Theta > (\alpha+\tau_2)\ (\tau_1-p_A)/\tau_2$ ....(17)

Thus for dumping to take place we have to have sufficient demand uncertainty in Country I.

Comment 2:

The lower bound on the foreign price in state A to make dumping an attractive option can be obtained as,

$$p_{A\,min} = [\tau_1 - \Theta\tau_2/(\alpha+\tau_2)]\ ;\ Q_e^* = 0 \text{ for all } p_A \le p_{A\,min}$$

Proposition 1:

The monopolist will choose a higher capacity level under trade compared to the optimal autarkic capacity.

Proof:

$$K^*_{trade} - K^*_{autarky} = [(\alpha + \tau_2)(p_A - \tau_1)/\tau_2 + \Theta]/\alpha > 0 \quad \text{by}(17)$$

III.2: Importing Country–the Analysis of Market Equilibrium

The market is perfectly competitive and the technology of the representative firm can be written as:

$$c(q,k) = c_1q + (c_2/2)(k-q)^2 + F_2 \qquad ....(2)$$

Here k is chosen ex ante on the basis expected price. Given k, since the adjustment cost is positive but finite, it is always profitable for the firm to adjust its production level optimally, knowing the ex post price. In each state of nature the representative firm is equating its marginal cost with the market price. Hence, upon a bad demand realization it will contract production from its precommitted capacity level to minimize loss whereas for a good demand realization actual output will exceed the capacity level to maximize profit.

When the firm produces at the capacity level, then its average variable cost equals marginal cost equals c_1. If expected price is greater than c_1, the larger the capacity, the larger will be the expected profit and, hence, the firm will choose $k=\infty$. If on the other hand, expected price is less than c_1, then the opposite is true and k=0. When expected price equals c_1 then the firm can choose any finite capacity level. Thus, the firm's capacity choice is given by

$$K = \begin{cases} \infty & \text{if } \bar{p} > c_1 \\ (0,\infty) & \text{if } \bar{p} = c_1 \\ 0 & \text{if } \bar{p} < c_1 \end{cases}$$

where $\bar{p} = \frac{1}{2}(p_A + p_B)$ is the expected price. By difinition, the industry's total capacity is nk, where n is the number of firms. Since the supply in the good (bad) state of nature is above (below) nk, and is positive and finite, it must be the case that nk>0 and is finite, implying

$$\bar{p} = c_1 \text{ or, } (p_A + p_B) = 2c_1 \qquad ....(18)$$

$$Q_A^* = 2(b-c_1)/a - 2\delta n/(an+2c_2) \qquad \text{.......(24a)}$$

$$Q_B^* = 2(b-c_1)/a + 2\delta n/(an+2c_2) \qquad \text{.......(24b)}$$

Profit in each state of nature can be calculated as follows:

$$\pi_i = (p_i - c_1)q_i - (c_2/2)[k-q_i]^2 - F_2 \quad i = A, B.$$

$$\text{or, } \pi_A = 2\delta^2 c_2/(an+2c_2)^2 - 4c_2\delta(b-c_1)/an(an+2c_2) - F_2 \qquad \text{.......(25a)}$$

$$\text{and } \pi_B = 2\delta^2 c_2/(an+2c_2)^2 + 4c_2\delta(b-c_1)/an(an+2c_2) - F_2 \qquad \text{.......(25b)}$$

Therefore, $E(\pi) = \frac{1}{2}(\pi_A + \pi_B) = 0$ gives the equlibrium number of firms as,

$$n^* = (\delta/a)(2C_2/F_2)^{1/2} - (2C_2/a) \qquad \text{.......(26)}$$

Proposition 2 :

$$n^* > 0 \text{ iff } \delta > (2C_2F_2)^{1/2}.$$

F_2 must not be too large to make production unprofitable at any positive level of output. Despite a constant ex ante cost c_1 expected total revenue exceeds expected variable cost. This arises from price variation induced by demand uncertainty and the firm's ability to adjust to the unanticipated change $(\tilde{c}_2)$. In the absence of fixed costs, each firm will have positive expected profit irrespective of the number of firms. The fixed cost F_2 determines the equilibrium number of firms through the 'zero expected profit' requirement.

When $\delta=0$ and $p_A=p_B=c_1$, then $n^*\leq 0$, since the firms can covver only variable cost and fixed cost are never recovered. Hence, the firm will not have any incentive to continue production. When $\delta>0$ and actual price is different from the expected price $\bar{p}$, the representative firm will always have the incentive to deviate from the capacity level as adjustment is not infinitely costly. In state A the price falls short of $\bar{p}$ by exactly the same amount by which price exceeds this expected level in state B. Symmetry of adjustment cost and the linearity of marginal cost together imply that, in both states, average variable cost is the same. As $q_A<½(q_A+q_B)$, the loss of profits in state A are more than compensated by the gain in state B and on average a firm earns more than its variable cost. To eliminate the excess expected profits, the fixed cost plays a crucial role.

Foreign Supply

Suppose in the bad state of nature, Country I dumps Q in Country II's market. To examine the effect of dumping on Country II's optimal capacity choice, output choice and market clearing prices, let us denote the equilibrium values in the 'no dumping' state by a superscript 'ND' and those after dumping by 'WD'.

If dumping is unanticipated, the total market supply would be Q_e^* above the autarkic level in state A, i.e.,

$$Q_{A,SR}^{WD} = Q_A^{ND} + Q_e^*.$$

The market clearing price in state A would be,

$$p_{A,SR}^{WD} = -(a/2)\left[2(b-c_1)/a - 2\delta n^*/(an^*+2c_2) + Q_e^*\right] + b - \delta$$

$$= p_A^{ND} - (a/2)\,Q_e^* \qquad \text{.......(27)}$$

Since $P_{A,SR}^{WD} < P_A^{ND}$ the actual loss in state A will be much higher than the anticipated loss. In the short-run, dumping lowers producer's profits and raises consumer surplus through lower prices. Hence, the overall welfare effect is ambiguous.

In the longer run, the capacity choice can be altered to bring expected profits back to zero. In this case, the volume of dumping will be determined simultaneously with four other variables in five equations. The system of equations is given as :

$$(2/a)\left(b-\delta-P_{A,LR}^{WD}\right) = (n/c_2)\left[p_{A,LR}^{WD} - c_1\right] + K_{WD} + Qe \quad \text{.......(28)}$$

$$(2/a)\left(b+\delta-P_{B,LR}^{WD}\right) = (n/c_2)\left[p_{B,LR}^{WD} - c_1\right] + K_{WD} \quad \text{.......(29)}$$

$$P_{A,LR}^{WD} + P_{B,LR}^{WD} = 2c_1 \quad \text{.......(18)}$$

$$Q_e = (2/\alpha)\left[(\alpha+\tau_2)\,(p^{WD}_{A,LR}-\tau_1)/\tau_2+\Theta\right] \qquad ....(15)$$

$$\tfrac{1}{2}\left(\pi^{WD}_{A,LR}+\pi^{WD}_{B,LR}\right)=0 \qquad .. ....(30)$$

and the unknowns are $\left[p^{WD}_{A,LR}\left(Q^{WD}_{A,LR}\right),\ p^{WD}_{B,LR}\left(Q^{WD}_{B,LR}\right),\ Q^*_e,\right.$ $\left.K_{WD},\ n^*_{WD}\right]$. From (28) and (15) we solve for $p^{WD}_{A,LR}$ to get,

$$p^{WD}_{A,LR}=\frac{2\,(b-\delta)\,c_2\alpha\tau_2+nc_1a\alpha\tau_2-Kac_2\alpha\tau_2+2(\alpha+\tau_2)\,\tau_1ac_2-2\Theta ac_2\tau_2}{(an+2c_2)\,\alpha\tau_2+(\alpha+\tau_2)\,ac_2} \qquad ....(31)$$

From (29) we get,

$$p^{WD}_{B,LR}=\frac{2\,c_2(b+\delta)+anc_1-ac_2K}{(an+2c_2)} \qquad ....(32)$$

Equations (31), (32) and (18) jointly yield,

$$K^*_{WD}=\frac{2\,(b-c_1)}{a}+\frac{(\alpha+\tau_2)[2c_2\delta+\phi(an+2c_2)]-\Theta\tau_2(an+2c_2)}{(an+2c_2)\alpha\tau_2+(\alpha+\tau_2)ac_2} \qquad ....(33)$$

where $\phi=(\tau_1-c_1)$.

Substituting K from (33) into (31) and (32) we get,

$$p^{WD}_{B,LR}=p^{ND}_{B}+\frac{ac_2(\alpha+\tau_2)\left[-\phi-2c_2\delta/(an+2c_2)\right]+ac_2\Theta\tau_2}{(an+2c_2)\alpha\tau_2+(\alpha+\tau_2)ac_2} \qquad ....(34)$$

and,

$$p^{WD}_{A,LR}=p^{ND}_{A}+\frac{ac_2(\alpha+\tau_2)\left[-\phi-2c_2\delta/(an+2c_2)\right]+ac_2\Theta\tau_2}{(an+2c_2)\alpha\tau_2+(\alpha+\tau_2)ac_2} \qquad ....(35)$$

using the necessary and sufficient condition for below-cost

Proposition 5 :

Compared with the 'no dumping' situation the equilibrium quantity demanded in the good (bad) state will fall (rise) in the presence of dumping.

Proposition 6 :

Compared with the 'no dumping' situation, in both states of nature the equilibrium volume of domestic production will fall in the presence of dumping.

The equilibrium number of firms can be obtained by solving equation (30). Here, the algebra is messy and intractable. So we will only try to provide some indications about the possible direction of change.

Since K=nk and k falls after dumping, either k or n or both will decrease. It can be argued that k can not increase. If k increases then the increase in 'loss' in state A will be overcompensated by the increase in 'gain' in state B and with same fixed costs(F_2) the representative firm will make positive expected profits. This is a contradiction to the breakeven condition. By the same reasoning it can be argued that k cannot remain unchanged. Hence, k has to fall. Whether n increases, decreases or remains unchanged will depend on the relative magnitude of the demand stochasticity parameter and cost adjustment parameter. It appears that none of these three possibilities can be ruled out (see figure 3), though it is more likely for n to fall. Appendix A pesents a more detailed analysis of the relationship.

The intuition behind the propositions derived above is fairly simple and may not depend on the particular form of the demand function. Diagram 4 shows that, whenever this particular specification of cost structure holds, all the aforesaid resultis can be observed even with a non-linear demand function[2]. In the 'no dumping' situation, $S(K_{ND})$ is the industry supply curve in both states of nature and with D_A and D_B as the respective state-specific demands, P^{ND}_A and P^{ND}_B would be the equilibrium prices. In the presence of dumping, and given the precomitted industry level capacity in the short run, state A supply curve will be $S^A(K_{ND})$ and state B supply curve will remain unchanged at $S(K_{ND})$, as there is dumping only in state A. The difference between $S(K_{ND})$

and $S^A(K_{ND})$ reflects the fact that the volume of dumping is a direct function of p_A. In the short run, the state A market clearing price would drop to $p^{WD}{}_{A,SR}$ which violates the viability condition [equation (18)] that $E(p)=c_1$. Hence, firms will revise their optimal capacity choice and at the industry level $K^*_{WD} \ngeq K^*_{ND}$ since, whenever $K^*_{WD} \geq K^*_{ND}$, $E(p) < c_1$. So, industry level capacity would be cut back. This will shift the supply curve inwards. This will affect both the good and bad state equilibrium prices, even when there is no dumping in state B. The intuition behind this result is relatively straight-forward. With dumping in state A, the actual loss based on prior capacity becomes greater than the expected loss. To make expected profits zero across different states of nature, the representative firm will cut back its optimal capacity. With a smaller capacity level, upon the realization of good demand, the domestic firms alone will not be able to meet the market demand at the good state 'no dumping' price, as there is no foreign supply in the good state. This excess demand will push up the equilibrium price in state B.

After the revision of industry level capacity, $S^B(K_{WD})$ and $S^A(K_{WD})$ wouldb the good and bad state supply schedule respectively, whereas D_A and D_B will remain unchanged. The p_A,min is the critical domestic price level such that when the domestic price in the bad state goes below this, dumping becomes an unattractive option to the exporting country's monopolist, given his demand and cost parameters. After dumping and the revision of the optimal capacity choice, in the long run $p^{WD}{}_{A,LR}$ and $p^{WD}{}_{B,LR}$ will be the equilibrium prices satisfying the viability condition $E(p)=c_1$. It should be noted here that, even with non-linear demand curves, all the above four propositions hold.

Section IV: Welfare Analysis of Dumping for the Importing Country

Let us now analyze the welfare effects of below-cost dumping. The producer's surplus is unaffected, as the firms' capacity levels are adjusted in equilibrium to achieve zero expected profit. Consumer surplus will increase in state A as price decreases, and will decrease in state B as here price increases.

Values of Θ and δ are restricted so that $\Theta \geq \delta$. The same holds for the values taken for α and a. Since we assume that country II has a relative cost advantage over country I, we require that $\tau_2 \geq c_2$. It has also been checked that all possible combination of parameter values satisfy the necessary and sufficient condition for c-dumping, given by (36). The simulation produced $3^7(=2187)$ values E(ΔCS) and all of them were negative. Regarding the direction of change the following observations are in order:

$dE(\Delta CS)/dc_2 < 0$, $dE(\Delta CS)/d\tau_2 < 0$, $dE(\Delta CS)/da >$),

$dE(\Delta CS)/d\phi > 0$, $dE(\Delta CS)/d\delta < 0$, $dE(\Delta CS)/d\Theta > 0$,

for all combination of values. Whereas, $dE(\Delta CS)/d\alpha \gtrless 0$.

From the above reported sign pattern, we observe that as the demand uncertainty and/or adjustment cost increase in either country, the expected consumer surplus falls. If it was negative to begin with, then high values of Θ, δ, τ_2, c_2 would tend to make it more negative. Table 3 presents some of the numerical values generated by the simulation exercise.

Section V: Technological Inflexibility, Distortion and Alternative Policy Measures

In this model the perfectly competitive domestic market structure requires the atomistic firms to break-even. Under the condition of demand uncertainties this can be ensured by equating expected price with expected marginal cost in both "pre" and "post" dumping situations. The presence of dumping reduces the price in state A, making an increase in the price in state B necessary. But that implies a cut back in the industry level capacity. As the capacity for the industry as a whole decreases, compared to the pre-dumping situation the cost of production and, hence, the supply price will increase for all positive output levels. That is why in figure 5 the domestic supply curve before dumping [$S(K_{ND})$] lies to the right of that after dumping [$S^B(K_{WD})$] . So, there is a production distortion caused by the technological inflexibility.

Again, consumer surplus will increase in the bad state of demand whereas it will decrease in the good state. In figure 5, the triangle ABC shows the gain in consumer surplus due to dumping in state A and triangle UVW represents the loss of consumer surplus in state B. Whenever ABC<UVW, dumping is welfare reducing.

Here we are implicity assuming the ability of the importing country to influence the price of its importables—which is simply the large country assumption. So, the lowering of the price in state A is a manifestation of the terms of trade distortion.

Suppose, the overall change in consumer surplus in the presence of offshore supply is negative. To eliminate this distortion either trade policy or domestic policy can be adopted. We will consider the relative efficiency of alternative policy instruments. Three trade policies and two domestic policies discussed here are import tariffs, quota restrictions, antidumping enforcements, production subsidy and consumption subsidy, respectively.

Imposition of import tariffs will cause imports to fall at all price levels in state A. After longrun optimal revision, the industry level capacity will not contract as much as it would if there were no tariffs. Thus tariff is generating (tariff) revenue and is positively contributing to the consumers' welfare in state B. The tariff revenue function is bell-shaped, which shows the amount of revenue generated to be zero at tariff levels zero and $(p_A - p_{A,min})$, where $p_{A,min}$ is the critical price of Country II which makes exports unattractive to the monopolist of Country I (confer figure 5). Here the optimum tariff will depend on two factors: overall change in consumer surplus across the different states of nature due to an increase in domestic industry-level capacity following a reduction in the volume of imports after the imposition of tariffs, and the magnitude of tariff revenue. The first effect will depend on a number of considerations, viz., domestic demand parameters, cost parameters, and so on. In the absence of exact specification of these functions, it is not possible to determine the optimum tariff level (t^*). However, it is highly likely that $t^* < (p_A - p_{A,min})$. When the quota is auctioned off competitively, then it will have the same effect as a tariff. So, an optimal quota is equivalent to an optimal tariff.

Antidumping measures will force the foreign monopolist to stop price discrimination. Here the monopolist's decision to "serve" or "not to serve" the foreign market will depend on the profit possibilities of these two alternatives. In the below cost dumping case, in fact, the exporter will be forced to charge prices not less than his long-run average cost of production (τ_1). As $\tau_1 > c_1$, by specification of our model, prohibition of dumping here will make the volume of foreign sale zero. However, while antidumping enforcement affects the domestic production distortion in the same way as does import tariff (quota), it forgoes the terms of trade benefits.

The government in Country II can offer state-contingent consumption subsidy to compensate for the loss of consumer surplus in state B. A production subsidy can produce the same kind of welfare effect. While a consumption subsidy will produce the effect through shifting the demand curve in state B outwards, a production subsidy will generate the same effect by shifting the supply curve rightwards. A third alternative policy prescription would be a combination of both production and consumption subsidies. Figure 6 explains that effect more clearly.

It is apparent from figure 6 that to compensate for the loss of consumer surplus in state B through the consumption subsidy alone, the after-subsidy demand curve has to be D_B'. To achieve the same objective through the production subsidy alone, the supply curve has to be $S(K_{ND})$. The dotted lines D_B'' and S_B'' represent a combination of consumption and production subsidies to ensure the same effect.

Purely domestic subsidy policy will put pressure on the government exchequer while all the commercial policies can achieve the same and without involving any direct expenditure on the part of the government. So, the first-best policy in this situation will be an optimum tariff coupled with a production/consumption subsidy.

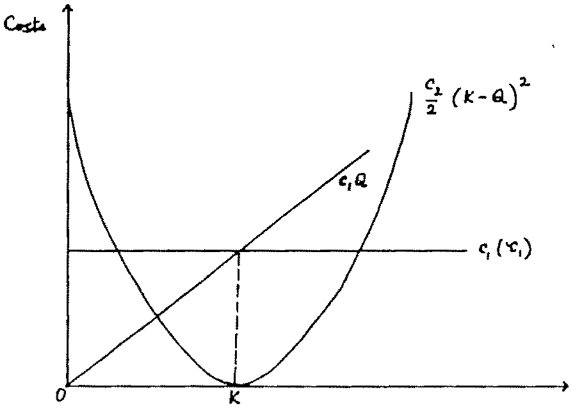
4.5 Summary and Conclusions

This paper shows that, when the product market does not adjust perfectly due to some technological inflexibilities, dumping is not unambiguously welfare improving. Dumping in the bad state of nature leads to a loss of consumer surplus in the good state as domestic production is cut back in each state of demand. These results have been established by using a linear cost function with quadratic adjustment costs and a linear demand function with additive randomness. To get a better feel of the analysis, a simulation exercise is carried out and the net welfare effect of below cost dumping on the importing country is calculated under different possible combinations of parameter values. The estimated net change in consumer surplus is negative in all cases.

Thus, the welfare effect of dumping in the importing country depends on the nature of consumer's preferences and the level of technology available to both the exporting and the importing countries. Hence, antidumping measures need not always be solely politically motivated; they might be justified on efficiency grounds.

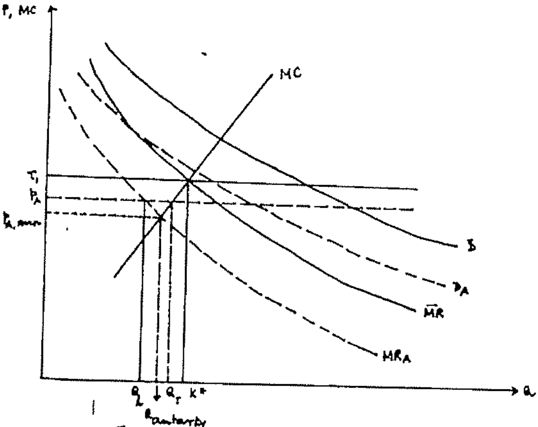
The possibility of dumping in the bad state leads to a decrease in domestic production of the good in question in both states of nature. This may cause unemployment of some factors in the import-competing sector, and a loss of income in excess of the net benefit (loss) to consumers consequent to the lower prices in state A. Since in reality, most of the time, the adjustment in factor markets is not instantaneous, the lowered volume of production in the import-competing sector will lead to unemployment and other related social losses, especially for relatively industry specific factors. Of course, this is an empirical question which should be studied at length by using appropriate econometric models and data.

Even when the income-generation aspect of dumping is assumed away, the policy analysis presented in the previous section shows that to cure the production and consumption distortions generated by technological inflexibilities, import tariffs coupled with consumption and/or production subsidization seems to be the best. The two features of our model producing these results are: (i) the major role played by the cyclical demand uncertainties in creating the distortion and (ii) the ability of the importing country to influence the price of its importables. While the first feature makes the welfare-effects of trade uncertain, the second one explains why trade policies are not necessarily dominated by the pure domestic policies. In fact, in the presence of large country advantage, trade policies are always more efficient in making full use of the "terms of trade" benefits.



The ex post Average Variable Cost and Average Total Cost for a given Level of Capacity

Fig.-1.



Optimal Output for the Exporting Country in the Bad State of Demand Before and After Trade

Fig.-2.

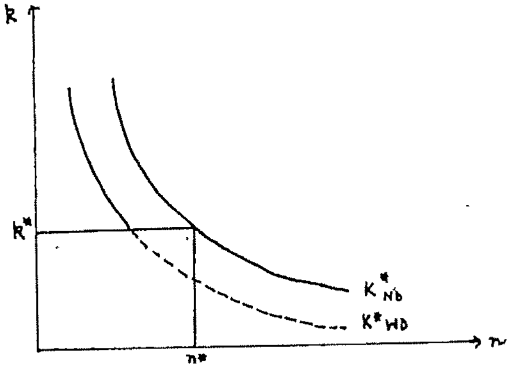


Figure 3: Firm-Level Capacity and the Equilibrium Number of firms in the Importing Country Before and After c-Dumping

Fig.-3.

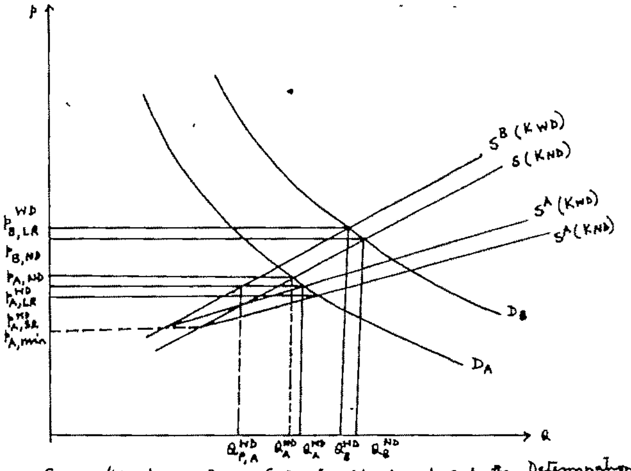
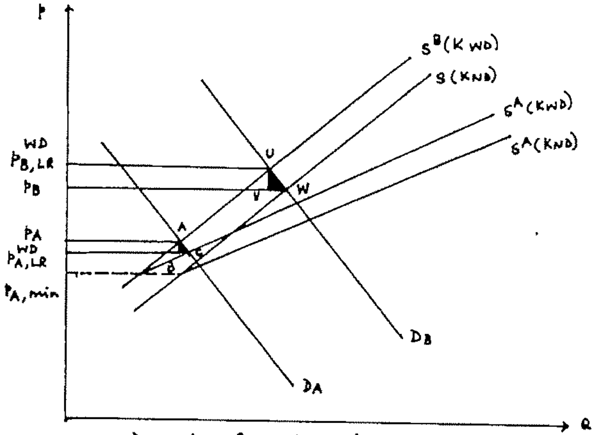


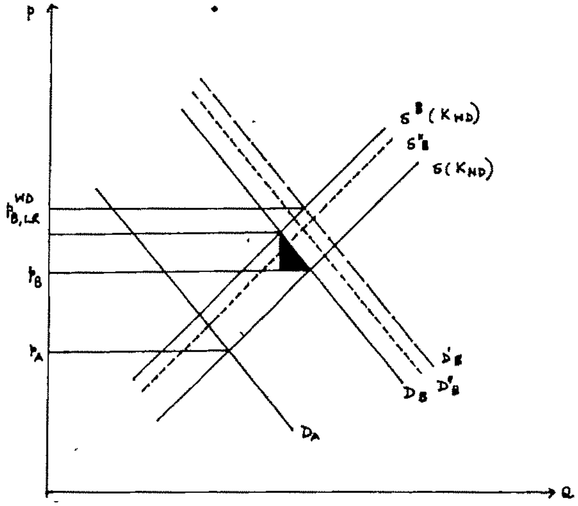
Figure 4: Long-Run Capacity Adjustment and the Determination of Equilibrium Price-Quantity over the Phases of Business Cycles in the Importing Country

Fig.-4.



c. Dumping Generated Distortions in Country II

Fig.-5.



Distortions and Domestic Subsidy in Country II During the Good State of Nature (State B)

Fig.-6.

Table 1 : Comparative Statics of (R–T) and M (signs only)

Denominator → Numerator ↓	$d\Theta$	$d\delta$	$d\alpha$	da	$d\tau_2$	dc_2	$d\phi$
d(R–T)	<0	>0	>0	?	?*	<0**	>0
dM	>0	<0	<0	>0	?*	>0**	<0

* Though the sign of both $d(R-T)/d\tau_2$ and $dM/d\tau_2$ are ambiguous, they are always opposite.

** $d(R-T)/dc_2 < 0$ and $dM/dc_2 > 0$, if $ac_2/\alpha\tau_2\,(an+ac_2) \le 0$.

Table 2 : Parameter Values Used in the Simulation Exercise*

Parameters → Values ↓	Θ	δ	α	a	τ_2	c_2	ϕ
Small	4.0	3.0	1.2	1.5	1.5	1.0	1.00
Medium	6.0	4.0	1.5	1.9	2.0	1.2	1.50
Large	8.0	5.0	1.7	2.	2.5	1.5	1.75

* Since market in the importing country is perfectly competitive n should be reasonably large. We have taken n=100.

Table 3 : Overall change in Consumer Surplus Following Below Cost Dumping

Θ	δ	α	ϕ	a	τ_2	c_2	ΔCS
4.0	3.0	1.2	1.00	1.5	1.50	1.00	– 0.1358
						1.20	– 0.1605
						1.50	– 0.1961
					2.00	1.00	– 0.1493
						1.20	– 0.1767
						1.50	– 0.2163
					2.50	1.00	– 0.1574
						1.20	– 0.1864
						1.50	– 0.2285
				1.9	1.50	1.00	– 0.1086

Table 3 : Overall change in Consumer Following Below Cost Dumping (*Contd.*)

Θ	δ	α	ϕ	a	τ_2	c_2	ΔCS
						1.20	– 0.1286
						1 50	– 0.1577
					2.00	1.00	– 0.1192
						1.20	– 0.1413
						1.50	– 0.1737
					2.50	1.00	– 0.1255
						1.20	– 0.1490
						1.50	– 0.1833
				2.2	1.50	1.00	– 0.0943
						1.20	– 0.1119
						1.50	– 0.1375
					2.00	1.00	– 0.1035
						1.20	– 0.1229
						1.50	– 0.1513
					2.50	1.00	– 0.1090
						1.20	– 0.1295
						1.50	– 0.1595
		1.7	1.50	1.5	1.50	1.00	– 0.0324
						1.20	– 0.0377
						1.50	– 0.0449
					2.00	1.00	– 0.0522
						1.20	– 0.0614
						1.50	– 0.0744
					2.50	1.00	– 0.0641
						1.20	– 0.0756
						1.50	– 0.0922
			1.75		1.50	1.00	– 0.0083
						1.20	– 0.0089
						1.50	– 0.0094
					2.00	1.00	– 0.0312
						1.20	– 0.0364
						1.50	– 0.0436

Appendix A : Equilibrium Number of Domestic Firms in the Importing Country Before and After Dumping

Equilibrium number of firms in the "no foreign supply" situation, in country II, is obtained by solving the zero (expected) profit equation for n, where state-specific profit functions are given by equations (25a) and (25b) When each firm in the industry is earning zero expected profit then the expected profit for the industry as a whole would also be zero. This implies,

$$P_A^{ND}Q_A^{ND} - c_1Q_A^{ND} - (c_1/2n)\ (K_{ND} - Q_A^{ND})^2 - nF_2 +$$

$$P_B^{ND}Q_B^{ND} - c_1Q_B^{ND} - (c_2/2n)\ (K_{ND} - Q_A^{ND})^2 - nF_2 = 0$$

$$\text{or, } (c_1 - p_A^{ND})\ (Q_B^{ND} - Q_A^{ND}) - 2F_2n - (c_2/2n)\left[2(K_{ND} - Q_A^{ND})^2\right] = 0 \qquad \text{.......(A.1)}$$

by using equations (23a), (23b), (24a), and (24b).

$$\text{or, } 2F_2n^2 - (c_1 - p_A^{ND})\ (Q_B^{ND} - Q_A^{ND})n + c_2(K_{ND} - Q_A^{ND})^2 = 0 \qquad \text{.......(A.2)}$$

This is a quadratic in n by solving which the equilibrium number of firms can be obtained as,

$$n^* = \frac{(c_1 - p_A^{ND})\ (Q_B^{ND} - Q_A^{ND}) \pm \left[(c_1 - p_A^{ND})^2\ (Q_B^{ND} - Q_A^{ND})^2 - 8F_2c_2(K_{ND} - Q_A^{ND})^2\right]^{1/2}}{4F_2} \qquad \text{.......(A.3)}$$

In the "with foreign supply" situation the corresponding condition to ensure zero (expected) industry profit will be :

$$2F_2n^2 - (c_1 - p_A^{WD})\ (Q_B^{WD} - Q_A^{WD})n + c_2(K_{WD} - Q_A^{WD})^2 = 0 \qquad \text{.......(A.4)}$$

and the solution to the quadratic gives the roots as

$$n^{**} = \frac{(c_1 - p_A^{WD})\ (Q_B^{WD} - Q_A^{WD}) \pm \left[(c_1 - p_A^{WD})^2\ (Q_B^{WD} - Q_A^{WD})^2 - 8F_2c_2(K_{WD} - Q_A^{WD})^2\right]^{1/2}}{4F_2} \qquad \text{.......(A.5)}$$

We have to compare n* with n**. If n* ≥ n**, then in the post dumping state, the total number of domestic firms will go down. Because, n* gives the equilibrium number of home firms alone, whereas n^{**} consists of both the domestic and the foreign suppliers. In fact, n^{**} would be the total number of suppliers when the foreign market is also competitive. If we can show that $n^{**}<n^{*}$, then for sure with a monopolist in the exporting country's market, the total number of firms (domestic) will be smaller after dumping. Since both (A.3) and (A.5) have the same denominator, comparison of the numerator will be sufficient for our purpose. Let us do it term by term.

Second term : equation (A.5)

$$(c_1 - p_A^{WD})^2 \, (Q_B^{WD} - Q_A^{WD})^2 \; - \; 8F_2c_2(K_{WD} - Q_A^{WD})^2$$

Now, $(K_{WD} - Q_A^{WD}) = (K_{ND} - Q_A^{WD}) - (an+4c_2)(G/c_2)$

by using equations (35), (37) and (38), where

$$G = \frac{c_2\left[(\alpha+\tau)\,\{-\phi - 2c_2\delta/(an+2c_2)\} + \Theta\tau_2\right]}{(an+2c_2)\alpha\tau_2 + (\alpha+\tau_2)ac_2} > 0$$

Hence, $(K_{WD} - Q_A^{WD}) < (K_{ND} - Q_A^{WD})$ ....(A.6)

Again, $(c_1 - p_A^{WD})^2 \, (Q_B^{WD} - Q_A^{WD})^2 \; - \; 8F_2c_2(K_{WD} - Q_A^{WD})^2$

$$= (c_1 - p_A^{ND})^2 \, (Q_B^{ND} - Q_A^{ND})^2 \; - \; 8F_2c_2(K_{ND} - Q_A^{ND})^2 +$$

$$\left[8G\{(c_1 - p_A^{ND}) + aG\}^2\{2G - (Q_B^{ND} - Q_A^{ND})\}\right] +$$

$$aG\,\{aG + 2\,(c_1 - p_A^{ND})\}\,(Q_B^{ND} - Q_A^{ND})^2 -$$

$$8F_2G\,(an+4c_2)\{c_2(an+4c_2)/G - 2\,(K_{ND} - Q_A^{ND})\} \qquad \text{.......(A.7)}$$

From (A.6) we get,

$$(K_{WD} - Q_A^{WD})^2 = (K_{ND} - Q_A^{ND})^2 +$$

$$(an+4c_2)(G/c_2)\{(an+4c_2)(G/c_2) - 2(K_{ND} - Q_A^{WD})\}$$

which implies,

$$(an+4c_2)(G/c_2)\{(an+4c_2)(G/c_2) - 2(K_{ND} - Q_A^{ND})\} < 0$$

or, $8F_2G\,(an+4c_2)\{(an+4c_2)(G/c_2) - 2(K_{ND} - Q_A^{ND})\} < 0$ as $8F_2c_2 > 0$

So, $2G > (Q_B^{ND} - Q_A^{ND}) \Rightarrow$

$$(c_1 - p_A^{WD})^2\ (Q_B^{WD} - Q_A^{WD})^2\ -\ 8F_2c_2(K_{WD} - Q_A^{WD})^2 >$$

$$(c_1 - p_A^{ND})^2\ (Q_B^{ND} - Q_A^{ND})^2\ -\ 8F_2c_2(K_{ND} - Q_A^{ND})^2$$

Sufficient Condition for $2G > (Q_B^{ND} - Q_A^{ND})$:

$$2G - (Q_B^{ND} - Q_A^{ND}) = 2G - 4\delta n / (an + 2c_2) = 2\left[G - 2\delta / \{a + (2c_2 / n)\}\right]$$

As $n \to \infty$, $\{a + (2c_2 / n)\} \to a$

Therefore, $2G > (Q_B^{ND} - Q_A^{ND})$ is implied by $G > (2\delta / a)$, for a sufficiently large number of firms ....(A.8)

First term : equation (A.5)

$$(c_1 - p_A^{WD})\ (Q_B^{WD} - Q_A^{WD}) = (c_1 - p_A^{ND})\ (Q_B^{ND} - Q_A^{ND}) - 4aG^2 +$$

$$aG\ (Q_B^{ND} - Q_A^{ND}) - 4G(c_1 - p_A^{ND})$$

by using equation (33), (34), (35), (37) and (38). Define,

$$H = 4aG^2 - aG\,(Q_B^{ND} - Q_A^{ND}) + 4G(c_1 - p_A^{ND})$$

$$= 4G\left[aG - \delta(an - 2c_2) / (an + 2c_2)\right]$$

by using (23a), (24a) and (24b).

As $n \to \infty$, $(an - 2c_2) / (an + 2c_2) \to 1$

Therefore, $H \to 4G(aG - \delta)$

Now, by condition (A.8) we have, $G > (2\delta / a)$

$\Rightarrow G > (\delta / a)$, or, $(aG - \delta) > 0$

Hence, $H > 0$ for sufficiently large n.

But, $H > 0 \Rightarrow (c_1 - p_A^{WD})\ (Q_B^{WD} - Q_A^{WD}) < (c_1 - p_A^{ND})\ (Q_B^{ND} - Q_A^{ND})$.

Therefore , when n* and n** are given by the smaller of the two sets of values yielded by equations (A.3) and (A.5), then n*>n**, i.e., the equlibrium number of firms decreases in the post dumping state, provided condition (A.8) holds.

Notes

This paper is a part of my Ph D. dissertation submitted to the Graduate School, University of Florida I am deeply indebted to Dr. Leonard Cheng for his supportive guidance and insightful comments Thanks also go to Drs. Steven Slutsky and Sanford Berg for their suggestions regarding the stylistic improvement of the presentation. of course, the responsibility of all the errors an omissions are solely mine

1 The comparative statics of equation (39) reveal that the volume of dumping is an increasing function of the adjustment cost (τ_2), demand stochasticity (Θ), and demand elasticity (which is inversely related to the slope parmeter 'α' of the inverse demand function) in Country I whereas it is a decreasing function of demand elasticity(inverse function of 'a'), demand stochasticity (δ) and relative cost (unit) advantage ($\varnothing$) of Country II. The sign with respect to Country II's adjustment cost (C_2) is indeterminate

2 Since the monopolist in the exporting country (Country I) is not a price taker, for him optimal price and quantity decisions will be taken simultaneously, given the demand and cost specifications. Here the cost function is specified precisely. However, because of this simultaneity, without any exact specification of the functional form of the demand relationship, an explicit solution for the choice variables is impossible to obtain An attempt is made to analyze the exporting country's problem in terms of figure 2, where deliberately the demand and marginal revenue functions are drawn as non-linear ones Moreover, for the importing country (Country II), where we explicitly solve for all the equilibrium values, the use of a generalized demand function made the algebra intractable. That is why we discussed the possible direction of results at a fairly intuitive level in terms of figure 4. It seems, however, that all the propositions derived in the text should hold good.

3. In fact, the magnitude also depends on the size of (R–T). But what we want to say here is that M cannot affect the direction of changes.

4 In our model the number of firms (n) is endogenously determined as a function of fixed cost (F). Here in the simulation exercise instead of specifying F we are specifying a particular value for 'n'.

References

Banerjee, S. (1989). "Dumping and Trade Policy Under the Condition of Demand Uncertainty and the Presence of Factor Specificity Together with Factor Price Inflexibility", *mimeo*, University of Florida, Gainesville 32611

Bernhardt, D. (1984). "Dumping, Adjustment Costs and Uncertainty", *Journal of Economic Dyanmics and Control*, 8, 349-370

Das, S. P. (1987) " Market Uncertainties and Cyclical Dumping", *mimeo*, Indiana University, Bloomington.

Hillman, A.L, and E.Katz. (1986). "Domestic uncertainty and Foreign Dumping", *Canadian Journal of Economics*, 19, 403-416.

Competitive South and Oligopolistic North: A Macro Model on Kaldorian Line

Prabirjit Sarkar*

Nicholas Kaldor in his Presidential address to the Royal Economic Society (July, 1976) presented an informal model of the world economy in the tradition of Keynes-Kalecki (Kaldor, 1976). He divided the world into primary and manufacturing sectors.

The primary producing sector produces at full capacity so that in response to demand-supply imbalance price adjusts. In the manufacturing sector, the greater part of production is concentrated in the hands of large corporations so that the prices are fixed by the producers themselves. A typical producer operates at less than full capacity. Industrial prices(in contrast to the prices of primary products) are not 'market clearing'.

The relative price or the terms of trade between the output of the two sectors move due to changes in the price of primary sector output. 'The more favourable are the terms of trade to agriculture and mining, the more current technological advance will be exploited through new investment, and the faster the growth of output' (ibifd, p.705). Thus movements in the terms of trade equilibrate the growth of saleable output of agriculture and mining with the growth of demand which in turn reflects the growth of the secondary (and tertiary) sectors.

This is the essence of a dual economy model Kaldor had in mind. On the Kaldorian line, the present paper develops a simple North-South growth model (Secton II).

Throughout the analysis, a fixed exchange rate regime has been assumed. The exchange rate has been set equal to unity. Existence of government budget has been ignored for simplicity. Technology is of the fixed coefficient type.

The North produces consumers' goods and machines under surplus capacity in an oligopolistic market structure. The South produces consumers' goods under a capacity constraint.

---

*This paper has been prepared to present at the 1993 Conference of the Bengal Economic Association *It has also been presented at the annual conference organised by the Dept of Economics, Jadavpur university.*

The capital (K)-output(Y) relationships in the two regions are given by the following equations:

$$Y_n = u_n q_n K_n, \ u_n < 1 \qquad \ldots\ldots.(1)$$

$$Y_s = u_s q_s K_s, \ u_s = 1 \qquad \ldots\ldots.(2)$$

where subscript n stands for the North and s for the South, Y is the level of output, K is the level of capital stock, u is the rate of capacity utilisation, q is the given productivity of capital.

Skipping the details of production and distribution in the South, it is assumed that a constant proportion of Southern income is saved and invested.

As the production level in the South is capacity constrained and saving is assumed to be identical to investment, the rate of growth in the South (g_x) is determined by the Harrod-Domar type equation :

$$g_x = q_s I_s / Y_s \qquad \ldots\ldots.(3)$$

or

$$g_x = q_s s_s / p \qquad \ldots\ldots.(4)$$

as $S_s = s_s Y_s$ and $I_s = S_s / p$.

Here I_s is the Southern demand for Northern machines for investment use (measured in terms of Northern goods) i.e. the level of physical investment on Northern machines, S_s is their total saving, s_s is the average and marginal propensity to save and p is the North-South terms of trade (price of Northern goods in terms of Southern goods).

Given the saving propensity (s_s) and the productivity of capital (q_s), the rate of growth in the South is inversely related to the North-South terms of trade as shown in the right hand part of Figure 1 (the g_x line).

Total consumption expenditure in the South is $(1-s_s)Y_s$. It has been assumed to be allocated between the Southern and Northern goods in fixed proportions, $f_s:(1-f_s)$.

Evidently, a constant proportion of Southern output is exported as shown by the following equation :

$$Ex_s = [1 - f_s(1 - s_s)]\, Y_s \qquad ....(5)$$

It then follows that Southern exports and output grow at the same rate. Hence the g_x line also represents the rate of growth of Southern exports. Demand-supply balance for Southern exports is ensured by price flexiblity under a competitive market structure.

There are two components of Southern demand for Northern goods– consumption, C_s and investment, I_s. Due to the assumption of saving-investment identity, the value of Southern exports is equal to the value of Southern imports as shown by the following equation :

$$Ex_s/p = C_s + I_s \qquad ....(6)$$

Here the both sides are measured in terms of Northern goods.

In the North, price is fixed by the producers. In the presence of surplus capacity, the price of Northern output (P_n) is set by applying a fixed mark up rate, t, on unit labour cost, L :

$$p_n = (1 + t)\, L \qquad ....(7)$$

where L is the product of fixed labour-output ratio, l_y and the money wage rate, W_n (assumed to be fixed by wage contracts).

Given the mark up rate, t, the real wage rate of the Northern workers is given in terms of Northern goods. Thus the mark up rate fixes the distribution of income between the Northern workers and the capitalists as envisaged in Kalecki(1971). The share of profit in Northern income m is given by the mark up rate as shown by the following equation :

$$m = t/(1+t) \qquad ....(8)$$

Unlike the saving-investment identity in the South, investment decision in the North is separate from the saving decision. The rate of capital accumulation (I_n/K_n) has been assumed to be a rising function of the rate of profit (r_n). In linear form, the investment function of the North is

$$I_n/K_n = a + b\, r_n \qquad ....(9)$$

where a and b are positive parameters.

Since the rate of profit, $r_n = mP_nY_n/P_nK_n = mq_nu_n$ (9) can be rewritten as

$I_n/K_n = a + iu_n$ ....(9)

where $i = mq_nb$, parametrically given.

Thus, in effect, the rate of capital accumulation is a rising function of capacity utilisation as shown by the G_I line in the left hand part of Figure 1.

It is assumed that the whole of profit income is saved–nothing is saved out of wage income. Given the share of profit, a constant fraction of Northern income is saved as in the South. This is shown by the following equation :

$S_n = mY_n$ ....(10)

Dividing both sides by the capital stock, K_n, it follows that the ratio of saving to capital stock is nothing but the rate of profit :

$S_n/K_n = mq_nu_n = r_n$ ....(10)

The Eq. (10) gives a proportional relationship between the ratio of saving to capital stock and the utilisation rate as shown by the G_s line in Figure 1.

Due to saving-investment identity in the South, the macro-balance in the North is given by the saving-investment equality in the North :

$S_n = I_n$ ....(11)

whence $S_n/K_n = I_n/K_n$

The equality (11) is shown by the point of intersection between the G_s and G_I lines. At this point the rate of growth of Northern output is g*. It is independent of the terms of trade,p.

Assuming that a constant proportion of northern wage income is spent on Southern goods for consumption purposes, it follows that g* is the rate of growth of demand for Southern exports. As this is independent of the terms of trade, the rate of growth of demand for Southern exports is shown by the horizontal line g_n in the right hand part of Figure 1.

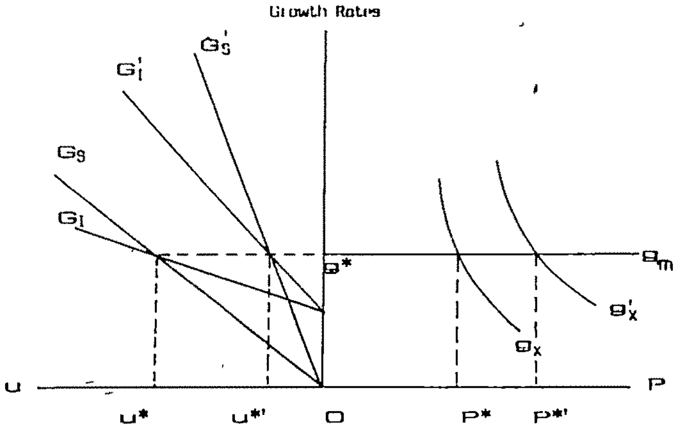


Fig.-1

The point of intersection between the g_x and g_n lines determines the equilibrium terms of trade, p*. The equilibrium rate of growth of the North and the South is g*.

Mathematically, the equilibrium solutions are given by the following equations :

$$u_n^* = a / mq_n(1-b) \qquad \text{.......(12)}$$

$$r_n^* = g^* = a/(1-b) \qquad \text{.......(13)}$$

$$p^* = (1-b)s_s q_s / a \qquad \text{.......(14)}$$

From the equilbrium solutions, (12) to (14), it appears that the following restrictions should be imposed on the investment function – Eq. (10) : a >0 and b <1.

Evidently, this framework casts serious doubts on the export-led growth strategy for the South as a whole. Stepping up the rate of growth of Southern exports requires a higher rate of growth of Southern output. Graphically it implies an upward shift in the g_x line to g_x'. This leads to a fall in the price of Southern goods in terms of Northern goods. But the growth rate of Northern demand for Southern exports given by the growth rate of the North (shown by the g_m line) is independent of the price ratio. Hence the South faces only a loss in the terms of trade (p* moves to p*').

Borrowing the words of Findlay (1981,p.455), it can be said that 'the model shows that while trade is the "engine of growth", in D.H. Robertson's famous phrase, the power that drives that engine is generated' by the 'growth rate of the North, the growth rate of the South adapting to the rate under the influence of changes in the terms of trade'. It then seems that the success stories of the NICs (Newly Industrialising Countries) cannot be repeated by the Southern countries as a whole.

it is also to be noted that a rise in productivity of capital in the South (q_s) due to, say, techincal progess, is fully reflected in the terms of trade – it turns against the South. But productivity changes in the North are not reflected in the movements in the terms of trade (see Eq. 14).

Thus as in Findlay (1980, 1981), the present framework lends support to the statement of Singer (1950) that 'technical progress in manufacturing industries showed in a rise in incomes while technical progress in the production of food and raw materials in underdeveloped countries showed in a fall of prices.'

Moreover, a glance at the Eqs. (13) and (14) shows that the mark up rate t or the monopoly power m does not figure in the equilibrium solutions for the terms of trade and the growth rate. That means a change in the monopoly power of the Northern capitalists to set their prices cannot influence the North-South terms of trade. Recently, Darity (1990) derived the same conclusion from a reformulation of Findlay (1980, 1981) model.

A rise in the mark up rate or the profit margin increases the share of profit in the Northern output. Firstly, this increases the rate of profit. As a result, the saving rate, S_n/K_n, rises. Graphically, this is shown by an upward rotation in the G_s line to G_s'. At the same time, it steps up the investment rate as shown by a similar movement in the G_I line to G_I'. Thus the dampening effect of increased saving on capacity utilisation is counteracted by an expansionary effect of increased rate of capital accumulation.

The net effect on capacity utilisation can be rigorously derived by differentiating (12) with respect to m :

$$du^*_n / dm = -a/(1-b)q_n m^2 < 0 \qquad ....(15)$$

From (15), it appears that capacity utilisation falls with the rise in the monopoly power of the Northern capitalists. It is this fall in capacity utilisation which nullifies the favourable effect of a rise in profit margin on the rate of profit. Hence in the new equilibrium, the rate of profit and the rate of growth remain unaffected.

Graphically, it implies that G_s and G_I lines rotate in a way so that u* falls to u*' but g* remains unaffected (Figure 1). At the same time, the North-South terms of trade also remains unaffected at its previous equilibrium value, p*

This result is in contrast to that found in the North–South model developed by Dutt (1988). He concluded that a rise in the monopoly power (and an income distribution favouring the Northern capitalists) slows down the rate of growth in the North and leads to a decline in the terms of trade of the South. However, it can be argued that this type of conclusion follows from his specification of the investment function.

Dutt (1988) took the rate of capital accumulation as a function of both the rate of profit and the utilisation rate:

$$I_n / K_n = a + br_n + cu_n \qquad \text{.......(16)}$$

$(a, b, c > 0)$

Using (11), the solutions for capacity utilisation and the rate of growth can be derived:

$$u_n^* = a / [mq_n(1-b) - c] \qquad \text{.......(17)}$$

and

$$g^* = mq_n u_n^* = amq_n / [mq_n(1-b) - c] \qquad \text{.......(18)}$$

Differentiating (18) with respect to m, it follows that

$$dg^*/dm = - acq_n / [mq_n(1-b) - c] \qquad \text{.......(19)}$$

Thus the specific investment function given in Eq. (16) yields the relationship· the higher the monopoly power and the share of profit, the lower is the rate of growth.

The objection to this type of investment function is that in a model where the mark up rate is parameterically given, endogenous variations in the rate of profit comes only through changes in the utilisation rate. Hence there is no justification of taking the rate of profit and the utilisation rate as two separate explanatory variables.

In fact, there is a lot of ambiguities in the relationship between distribution and growth as shown by Blecker(1989), Bhaduri and Marglin (1990) and Sarkar (1992; 1993a,b).

III

The Kaldorian model of the world economy takes us very close to the conclusions drawn in the neoclassical model of Findlay (1980 & 1981) and its reformulation by Darity (1990). In particular, as in Findlay (1980, 1981), our model shows that the South as a whole cannot speed up its long-term growth rate through export promotion; rather it faces a terms of trade loss. Next, like Darity (1990), this paper questions the theoretical validity of the explanation of the Prebisch-Singer hypothesis of secular decline in the terms of trade of the South in terms of the monopoly power of the Northern capitalists. Factors explaining the 'fact'[1] of secular decline in the terms of trade of the South vis-a-vis the North are more complex than a mere rise in the mark up rate in the Northern manufacturing sector.

NOTES

1 See Thirlwall *et. al*, 1985; Sarkar, 1986, Sarkar-Singer, 1991 and Sapsford *et. al*, 1992 for proving the statistical validity of the Prebisch-Singer doctrine.

REFERENCES

Bhaduri, A and Marglin, M., (1990), "Unemployment and the real wage: the economic basis for contesting political ideologies", *Cambridge Journal of Economics*, vol 14, pp. 375–393

Blecker, R A , (1989), "International competition, income distribution and economic growth", *Cambridge Journal of Economics*, vol 13, pp 395–412.

Darity W, Jr , (1990), "The fundamental determinants of the terms of trade reconsidered long-run and long-period equilibrium", *American Economic Review*, vol 80, pp 816–27

Dutt, A K (1988), "Monopoly power and uneven development . Baran revisited", *Journal of Development Studies*, vol 24, pp. 161–76.

Findlay, R , (1980), "The terms of trade and equilibrium growth in the world economy", *American Economic Review*, vol. 70, pp 291–99

---------- (1981), "The fundamental determinants of the terms of trade", in *The World Economic Order* (eds.), S Grassman and E Lundberg, London Macmillan, pp 425–57

Kaldor, N. (1976), "Inflation and recession in the world economy", *The Economic Journal*, vol 86, pp 703–14

Kalecki, M., (1971), *Selected Essays in the Dynamics of the Capitalist Economy*, Cambridge: CUP.

Sapsford, D , Sarkar, P , and Singer, H W , (1992), "The Prebisch-Singer terms of trade controversy revisited" *Journal of International Development*, vol. 4, pp. 315–332.

Sarkar, Prabirjit, (1986), "The Singer-Prebisch Hypothesis . a statistical evaluation", *Cambridge Journal of Economics*, vol. 10, pp. 355–371.

---------- (1992), "Industrial growth and income inequality . an examination of 'stagnationism' with special reference to India", *Journal of Quantitative Economics*, vol. 8, pp. 125–138.

---------- (1993a), "Effective demand and income distribution in the context of agriculture-industry demand linkage : a two-sector macroeconomic framework", *Journal of Macroeconomics*, vol. 15.

---------- (1993b), "Distribution and growth : a critical note on 'stagnationism'", *Review of Radical Political Economics*, vol. 25.

---------- and Singer, H. W , (1991), "Manufactured exports of developing countries and their terms of trade since 1965", *World Development*, vol. 19, pp. 333–340

Singer, H. W., (1950), "The distribution of gains between investing and borrowing countries", *American Economic Review*, vol 40, pp 473–485.

Thirlwall, A.P , (1986), "A general model of growth and development on Kaldorian lines", *Oxford Economic Papers*, vol 38, pp 199–219

International Trade and Uneven Development: Implications of Some North-South Models

Amitava Krishna Dutt*

Introduction

The phenonenon of uneven development–in which success breeds succes while at the same time failure breeds failure-appears to be common enough. One finds frequent examples of virtuous and vicious circles, and when the two cycles are found side by side we observe uneven development. In economics there are examples at the level of individuals and firms, and at the level of aggregates such as industries, regions, nations, and groups of nations. In this paper I will be concerned with development between nations, and a simple way to examine this is to consider groups of nations–one group that we may call the North and the other South, representing developed and underdeveloped regions, respectively.

To start with, one can ask if this type of uneven development in fact occurs. Historically, there is no doubt that it has occured. Using the economists' pet measure-per capita income,for instance, between 1757 and 1947, British per capita income increased at least tenfold, while Indian per capita income at best did not increase, and probably fell (see Maddison,1971). In more recent years, however, there may be some evidence against this phenomenon. During 1965-90, high-income OECD economies (which we may identify with the North) experienced average annual GNP per capita growth of 2.4 per cent, compared to 2.9 for low-and medium-income countries. This may be taken to imply that development has tended to narrow North-South differences. One may perhaps echo the title of Rostow's (1980) book and suggest that the more relevant question is, Why the poor get richer and the rich slow down. But a closer look presents a different picture.

---

*An earlier version of this paper was presented at the University of Massachusetts at Amherst, and at the Second Annual Analytical Political Economy Conference at the American University, Washington D.C. Iam grateful to the participants, especially J. Mohan Rao, and Peter Skott, and to jose Cordero, for comments.

The growth rate has been 0.2 for Sub-Saharan Africa, 1.9 for South Asia, and 1.8 per cent for Latin America and the Caribbean (regions we may identify as comprising the South), all lower than the Northern rate. 'Rostow's title is a little misleading, as a reading of his paper with the title shows that 'poor' really refer to 'middle income countries'. Romer (1986) has also argued that growth rates seem to increase with the level of development historically and in more recent periods. Thus international uneven development appears to be an important phenomenon which is still a highly relevant one. Of course, this does not mean that there have not been important examples of poorer countries which have broken out of this phenomenon – the East Asian NIC's are the ones most commonly mentioned in this regard.

Let us suppose that uneven development in our sense is an observed fact. The question then arises: Is it part of the same phenonon – are vicious and virtuous cycles related? If they are not so interrelated there is no overwhelming reason for studying the two types of development simultaneously. To examine this issue, one needs to look beyond stylized facts and employ a theoretical approach to the question.[2] Two important theoretical approaches which still dominate the study of international economics and development are the neoclassical Heckscher-Ohlin-Samuelson approach and the Marxian/Dependency approach. These approaches, of course, are well known, and only a few points are worth raising here. The former approach is purely theoretical, generally using an optimizing framework and assumptions of perfect competition. The presumption here is – apart from exceptional cases of distortions – countries benefit from trade. The second approach has generally emphasized historical details, been informal, and discussed a variety of mechanisms which lead to surplus transfers between nations and other types of losses, generally implying uneven development. Both approaches arguably have important shortcomings. The first typically ignores the role of structures, getting bogged down in optimization exercises, and seems to pay no attention to history. The second is often vague (concepts such as surplus and surplus transfer are often not clearly formulated), and becomes teleological: uneven development is an iron law; the response to NIC development has sometimes (see Bienefeld, 1981), in this approach, been that the NICs have not developed, but experienced 'dependent' development, no development at all!

Recently a new literature has emerged which has tried to directly examine the question of North-South trade relations. They have borrowed from both perspectives noted above. They have produced a host of formal models, but which have paid attention to structures, generally departed from the fetish of optimization (though this is very hard for most mainstream economists[1]), and tried to incorporate some of the important mechanisms highlighted in the Marxist/Dependency/Structuralist uneven development literature. The aproach essentially looks at macroeconomic accounting relations involving domestic output use and distribution , international trade, and sometimes, capital flows, filling in this skeleton with appropriate behavioural, institutional and other relations, and examining the consequences of such assumptions for the path or the world economy.

In this paper my purpose is to discuss the usefulness of these models in relation to three important questions that the issue of uneven development raises: first, how did uneven development originate? Second, how is it propagated? Third, how, if at all does it end? I will present three simple models as illustrations of how to analyze each question, and show how exactly they must be related to the study of history and institutions for a proper understanding of the phenomenon of uneven development.

The origins of uneven development

How did uneven development orginate? I would argue that it is nost fruitful to study the history of particular interactions between sets of regions, because it is not at all clear that all such interactions occurred in the same way. Two interactions have been studied in the literature. First, there is the phenomenon of the Atlantic slave trade, and second, the phenomenon of deindustrialization due to colonialization in India. Darity (1982) and other writers have drawn on the insights of Rodney and others to analyze the former case. I will report on some issues raised by the latter.

In the seventeenth century the Indian subcontinent (henceforth : India) was not significantly less developed than Britain. But, as noted above between 1757, when Britain began colonizing India, and 1947, when India gained independence (and British growth had already slowed down), the ratio between British and Indian per capita income increased at least tenfold, and several countries initially less developed than India grew rapidly while India stagnated. No clear answer to the question of why this uneven development originated has emerged, in part because of the paucity of data, but also because of the heat and smoke generated by the debates on the impact of colonialism, which show no sign of abating (see Charlesworth, 1982). India's relative stagnation is explained in the 'nationalist' view by colonialism, and in the 'imperialist' view by the socio-economic structure established during the ancient Hindu period.

To explain this, it seems appropriate to start with a model which starts with an identical structure for the two regions, and then examines how the pattern of specialization assumed in the models noted above came about. A model of this type, originally developed by Krugman (1981) and further developed in Dutt (1986), shows how two otherwise identical regions, can evolve in such a way that if one region starts with a slight advantage over the other, it will keep forging ahead. This model belongs to a class of models which stress the inportantce of increasing returns in manufacturing and give an important place to history (Krugman, 1991).

Consider a model of a two-region world economy given by the following equations :

$$M_i = aK_i,\ L_{mi} = M_i / b_{mi},\ L_{ai} = L - L_{mi},\ A_i = b_a L_{ai} \qquad \text{.......(1)}$$

where M_i is manufacturing output, K_i capital (employed only in manufacturing), L_{ji} labor force (in sector j for m and a, denoting manufacturing and agricultural sectors), b_{mi} the productivity of labor in manufacturing, all in the two regions 1 and 2, denoted by subscript i, and where a is the output-capital ratio in manufacturing, b_a the productivity of labor in agriculture, and L the total labor force, assumed to be the same in each region. Assuming that each region produces both goods, and labor mobility within each region equalizes the wage between sectors in each region, the wage in each region is given by b. The rate of profit in manufacturing (where the manufactured good is the capital good) is thus given by

$$r_i = [1 - b_a / (b_{mi}\, p_i)]a \qquad \text{......(2)}$$

where p_i is the relative price of the manufactured good in region i. Profit-recipients in each region save a fraction s of their income, and a fraction α of all consumption expenditure is spent on the manufactured good, the rest being spent on the agricultural good. All saving in each region is invested within that region, so that, assuming away depreciation,

$$dK_i / dt = sr_i \qquad \text{.......(3)}$$

Finally, as capital is accumulated in each country, due to increasing returns and learning effects, labor productivity rises in the manufacturing sector, and this is captured by assuming that

$$b_{mi} = \beta K_i^{\tau} \qquad \text{.......(4)}$$

where β and τ are positive parameters, and τ<1 to reflect diminishing retuns to learning. Not that in this model all differences in the two regions are attributable to differences in K_i , and all other possible differences have been suppressed to ensure identity of structure.

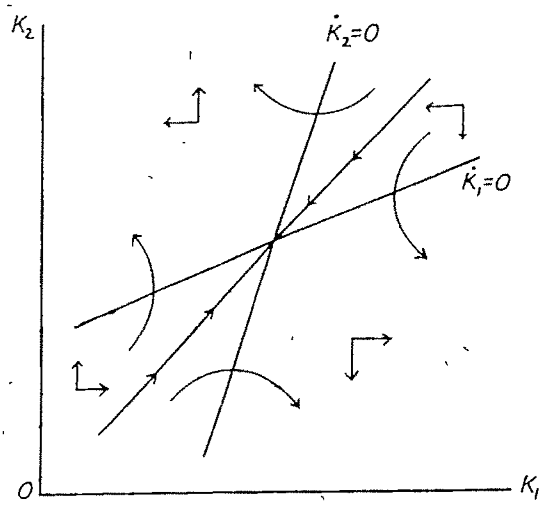


Fig.-1

If the two regions do not trade (factor movements between the regions is assumed away here) with each other, starting from any low levels of capital stock, they will accumulate capital over time, reaching a stationary state with $K_i = (\alpha \beta L / a)^{\frac{1}{1-\tau}}$. If, on the other hand, the two regions trade freely with each other, so that the p are equalized, then depending on which region has a slightly higher K_i, that region will experience steadily increasing K_i and the other region will (eventually) experience a decline in its K_i (although initially there may be no absolute decline, and even a modest relative change in the ratio of the stocks of capital). The region with the higher K_i will export the manufactured good, and as it experiences a greater growth in manufactured capital stock over time, it will experience a faster rate of technological change in manufacturing, earn a higher profit rate, and save and invest at a higher rate. Eventually, it will become completely specialized in the production of manufactured goods and the other region in the production of primary goods.[3] The dynamics of a possible outcome is shown in Fig. 1, where the $\dot{K}_i = 0$ line shows combinations of K_1 and K_2 which set $dK_i/dt = 0$. This model belongs to a general class of models which stress the importance of increasing returns and learning in manufacturing, which involves *processing*, and demonstrate how small changes in history can have large effects (see Arthur, 1990).

The model just described is obviously too simple to provide definitive answers to the question of the origins of uneven development for the case under discussion. India and Britain were parts of a world economy, which violates the two-region assumption of the model, and the structures of the two economies when they came into contact in the eighteenth century were not identical, as assumed in the model. But since Britain was India's major trading partner in the nineteenth century (Chaudhuri, 1982), and since the perfect symmetry assumption is not required of the validity of the result on the implications of the specialization, we may explore the consequences of taking the two regions of the model to represent India and the rest of the world dominated by Britain.[4]

India's pattern of trade changed drastically after a few decades of colonial rule (see Raychaudhuri, 1968 and Chaudhuri, 1982). When the British began their colonization, India was an exporter of industrial goods and an importer of primary and intermediate goods. Before 1800, India was the major supplier of cotton and silk textiles (fine clothes as well as everyday wear for the masses) in international markets, including Europe; Indian textiles were considerably cheaper than British woolens because of lower wages and technical advantages. In addition to these, which constituted the overwhelming bulk of India's exports, she exported raw silk, sugar and saltpetre, which do not fit the description of primary agricultural products; even indigo involved some amount of processing. The main imports were foodstuffs such as coffee, tea, sugar and spices, and luxury items such as wine and horses, and a considerable amount of precious metals (given India's chronic merchandise export surplus). Except for an insignificant amount of luxury goods, and occasionaly, cannon, no manufactured metal products were imported before the nineteenth century. After the first few decades on the nineteenth century, India became primarily an exporter of agricultural products, including raw cotton and jute, tea and coffee, opium, indigo, oil-seeds and foodgrains, and an importer of manufactures such as cotton yarns and cloth , metals and machinery. The case of textiles is most striking: Bengal handicraft manufactures, so dominant in the 1770s that an European observer wrote that their incomparable quality guaranteed that their demand would never falter due to competition, were by the first three decades of the nineteenth century almost completely eliminated from international markets. Between 1814 and 1835 exports from India to England of cotton goods fell from 1,266,608 pieces to 306,086, while imports rose from 818,208 yards to 51,777,277; a similar fate befell Indian exports to other markets (Dutt, 1903). By the turn of the century imported piece goods supplied about 60 per cent of Indian cloth consumption, a proportion which was probably higher earlier in the century (Maddison, 1971).

An examination of some features of foreign rule provide some plausible explanations for this shift in India's pattern of specialization, which have been carefully documented by nationalist historians (Dutt, 1901, 1903, Basu, 1935; see also Chandra, 1966). First, Indian handicraft industries were adversely affected by the decline of Indian royal courts, which had been important buyers of quality products and promoters of factories making to their order. Second, the same industries were dislocated by the exploitation by the English East India Company after it gained important trading privileges in Bengal, as its traders imposed arbitrary prices and other conditions, and subjected the artisans to flogging, imprisonment and worse (the cutting off of the thumbs of winders of raw silk has been documented). Third, while internal tariffs and transit duties restricted trade in Indian goods, British goods were granted exemptions. Fourth, while import duties were imposed in Britain on imports of Indian manufactures (in 1812 calicoes were taxed at about 72 per cent, and import taxes on other goods ranged from 100 to 600 per cent), British goods were allowed into India at no or low duties of 2.5 per cent.

These features, by worsening the competitive position of Indian handicrafts even slightly, could have led to the cumulative process of uneven development and structural change analyzed by our model, as hinted by early nationalist writers like Ranade and Joshi (see Chandra, 1966).

Another relevant consequence of British rule is the so-called 'drain' of wealth from India to Britain, initially as a result of direct plunder and looting, and later through the payment from Indian revenues administrative charges incurred in India and elsewhere in the empire, interest payment on debt, and remittance on retained profits. Various inconsistencies of the early presentation of this theory by Naoroji have been pointed out, and it has been argued that if one takes into account the fact that India received a great deal in return for the outflow of capital (railway construction, for instance), the drain becomes a negligible proportion of Indian income (see Charlesworth, 1982). Our model, however, implies that a relatively small drain of capital during the early stages of interaction between the two regions can explain the shift in the pattern of specialization, and if we extend the model to allow for unrequited capital flows through time, this drain exacerbates the tendency toward uneven development due to deindustrialization (a connection that was noted by Naoroji and others). It is relevant to note here, first, that the extent of the drain during the early period of foreign rule is less controversial than that of the later period and second, that the development of railways had the effect of reducing the effective protection of Indian industries after uneven development was already under way, and therefore exacerbated it.

Three comments addressing some possible objections to our analysis are in order at this stage.

First, the analysis may appear to be contradicted by claims made by 'imperialist' writers and their modern counterparts that British rule had some favourable effects on the rate of Indian economic growth through the establishment of law and order, the development of railways, and the spread of commercialization in agriculture; moreover, theory and empirical evidence is marshalled to argue that India experienced no absolute decline in industrial employment. However, our interpretation of the model is quite consistent with temporary increases in per capita output and modest increases in the industrial labor force (if it is extended to allow for population growth). Our model predicts that, other things constant, there would eventually be a stagnation in per capita income growth, and a decline in the share of the population dependent on industry (that this may have been slight can be explained by the relative self-sufficiency – due to inadequate transportation – of many village communities), and available evidence seems to corroborate these tendencies.

Second, it must be conceded that the kinds of events described above may not have been responsible for the reversal in India's pattern of specializartion, even in terms of our model. It has been claimed, instead, that technological change in British industry was responsible for the shift, and that this was independent of British rule in India. It is also suggested, to buttress this argument, that the reduction in import duties in Britain on Indian goods in the 1820s and 1830s did not revive Indian cotton good exports to Britain. However, against this it can be argued that had India been politically autonomous, she could have provided tariff and other forms of support to her industries. Moreover, even in 1813, witness after witness in the Select Committee of the House of Lords testified that free Indian textile imports (of both finer and coarser varieties) would damage British industry (Basu, 1935), and Wilson (1845, p. 538-9), based on this testimony, calculated that cotton and silk goods manufactured in India could be sold for a profit in Britain at a price 50 to 60 per cent lower than those manufactured in Britain. It may be speculated that the subsequent sharp fall (almost 60 per cent between 1819-21 to 1829-31 for cotton piecegoods) in British costs was due to learning resulting from the expansion of British exports to colonies such as India, and only after this occurred could tariffs be reduced without jeopardizing British industry.

Finally, it may be argued that our model takes an overly mechanical view of technical change and industrial growth: that it fails to distinguish between the handicraft industry of India and factory industry of Britain (with only the latter having a potential for scale economies and technical change), and that it fails to take into account the potential for industrial growth in an economy which is initially an exporter of primary products. Against the former claim it must be argued that conventional wisdom seems to overstress the distinction between factory and craft production and the technological advantages of the former (see Sabel and Zeitlin, 1985): in Europe and the US, craft production had all the technological dynamism of mass production. The role of small-scale handicrafts in the industrialization process in Meiji Japan is also well known. There is also evidence that the Indian cotton industry was flexible in expanding its production and adjusting its product to European tastes and adopting foreign methods (the shipbuilding industry also showing a capacity for imitative innovation); and the fact that it remained highly labour intensive and did not mechanize can be explained in terms of low wages, expanding markets, and the absence of competition (Raychaudhuri, 1982). This is not to argue that Indian craft industry was as sophisticated as the western cases of flexible specialization,or that India was on the brink of an industrial revolution before the British invasion. What is being argued, however, is that deindustrialization reduced by far (especially by hitting hardest the finer end of the textile industry) the chances that the Indian economy could experience industrial development based on its handicraft industry (perhaps by drawing skilled labour and entrepreneurship from it, or through technical change in the craft shops themselves), given fortuitous circumstances. On the latter point it must be conceded that our model overstresses the knife-edgedness of the uneven development process, ignoring various channels by which late-comers can industrialize. But as is well documented (see Dutt, 1901-3, Basu, 1935), India's foreign government, by obstructing Indian industrialization, or at least failing to systematically support it (as admitted even by modern critics of early nationalist views), arguably blocked these channels. A detailed examination of this history is required to sustain this point, but that is beyond the scope of this discussion.

One final comment: why is the North able to dominate the South? This is a fascinating question on the causes of Northern military superiority. Findlay (1992) has speculated on this issue and argued that it is because of the North's dependence on trade to obtain necessary goods – an inversion of the standard dependency argument! – that it was led to combine ships with cannon. The South also had ships and cannons, but did not combine them to turn trade to their advantage.

The perpetuation of uneven development

Assume that the North and the South are now differentiated. One difference is directly related to the simple model discussed in the previous section. According to that model, eventually one region-the North - will produce only manufactured goods, and the other region - the South - will produce only primary products. Even if this extreme result is not realistic, it can be taken to be an approximation to the difference in production structures of rich and poor nations. A second difference, which may emerge as a result of the uneven development process which has already started, is that the two regions may begin to show additional differences due to the differences in their growth experience. For instance, the arrested development in the South may eventually imply that there will be a large pool of unemployed labor with their wage tied to a subsistence level, while the more rapid development in the North will imply growth with its labour fully employed. Without examining how such different structures precisely came about, we can examine the implications of such structural defferences (both those just mentioned) between the two regions for the gap between the two regions. Two issues need to be looked into before such an analysis can be conducted.

First, if structural differences between the North and the South of the second type noted above do emerge, one needs to characterize in a precise manner such differences in structure. Here a variety of alternative models have been developed, specifying the structures of the two regions. For the South the standard assumption, consistent with the possibility noted in the previous paragraph, is to assume that there is unemployed labour available at a fixed subsistence level. For the North, however, different assumptions have been made in the literature. Findlay (1979,

1980) has assumed that growth occurs with fully employed labor as in neoclassical models of the Solow type, Taylor (1981, 1983) assumes that effective demand constrains Northern growth, as in Keynes-Kalecki models, while Vines (1984) and Molana and Vines (1989) assume that the North also has a fixed real wage, following some models of Kaldor.

Second, one can examine a variety of *mechanisms* which may possibly propagate uneven development. A large number of mechanisms have been suggested in the literature of North-South models, attempting to formalize the informal discussion on uneven development. Some have stressed the importance of increasing returns, continuing in the vein of the model of the previous discussion (see Dutt, 1986, Conway and Darity, 1990). Others have focussed on the role of Engel effects and changes in demand composition in the North and South, formalizing the idea that a larger gap between Northern and Southern development will imply shifts in demand toward Northern goods (see Dutt, 1988). Yet others have incorporated international debt issues in which uneven development can result from increases in the interest rate that the South pays to the North as the former's debt-GNP ratio rises. One important result found in these models is that while uneven development may emerge in these models, it is not always inevitable. Thus these models not only show the existence of what Myrdal (1958) called backwash effects which propagate uneven development and spread effects which reduce uneven development, but develop conditions which show whether it is more likely that uneven development will in fact occur, that is, the backwash effects will dominate the spread effects. Moreover, they have implications for the relation between international development and terms of trade changes of the South relative to the North, which has been widely discussed in the literature (see Singer, 1987).

To illustrate some of these issues, let us consider one model formalizing a persistent theme in the analysis of the economic relations between North and South which relates the increasing gap between the regions caused by technological change in the North which replaces imported Southern raw materials by synthetic raw materials and products (Nurkse, 1959, Cairncross, 1960, Singer, 1987).[5]

Consider two regions, the North and the South, each producing a single good. The Northern good is used as a consumption and a capital or investment good in both regions, while the Southern good is used as a consumption good in the South or as an intermediate good in the Norh (it is not traded as a consumption good).[6]

In the North firms produce according to the fixed-coefficients production relation

$$X_n = \min\ [a_n K_n,\ L_n / b_n,\ M / c] \qquad ....(5)$$

where X_i denotes output, a_i the output-capital ratio, b_i the labour-output ratio, K_i the stock of capital, and L_i the amount of labour empolyed, all in region i, where n denotes the North, M is the amount of raw material used, and c the per-unit raw material requirement. While Northern production requires capital, labour and intermediate good, in the South production uses only labour and land. The amount of land is taken to be given, but the effective amount of land can be indefinitely increased by using more capital. It is assumed that production technology can be represented by

$$X_s = \min\ [K_s a_s,\ L_s / b_s] \qquad ....(6)$$

where subscript s refers to the South.

In both regions labour is assumed to be in abundant supply at a fixed real wage denoted by V_i, in terms of goods consumed by workers. Denoting the money wage in i by W_i and the money price of the good produced in region i by P_i,[7] in the North this implies that

$$W_n / P_n = V_n \qquad ....(7)$$

while in the South this implies that

$$W_s / (P_s^m P_n^{1-m}) = V_s \qquad ....(8)$$

where we assume that a fraction m of Southern consumption expenditure is devoted to the Southern good, and the rest to the Northern good.

The markets for both goods are competitively organized and prices vary to clear the markets with capital fully utilized in each region. Price equations are therefore given by

$$P_n = (r_n / a_n)\ P_n + b_n W_n + cP \qquad ....(9)$$

and

$$P_s = (r_s / a_s) P_n + b_s W_s \qquad ....(10)$$

where r_i is the profit rate in region i defined as the rate of retun, per unit of capital, after payment of wages and raw materials (the latter only in the North).

Workers who earn wages, consume their entire income in both regions, and profit-recipients or capitalists, save a fraction s_i of their income in each region and always invest all their savings. Remembering that a fraction m of Southern consumption expenditure goes to the Southern good, the market-clearing conditions for the two goods is given by

$$P_n X_n = (1-m)\left[W_s b_s X_s + (1-s_s) r_s P_n K_s\right] + W_n b_n X_n$$

$$+ (1-s_n) r_n P_n K_n + P_n (I_n + I_s) \qquad ....(11)$$

and

$$P_s X_s = m\{W_s b_s X_s (1-s_s) r_s P_n K_s\} + cP_s X_n \qquad ....(12)$$

where I_i denotes investment in region i.

Finally, we abstract from inter-regional capital and labour flows, and assume that trade between the two regions is balanced, so that

$$cP_s X_n = (1-m)\{W_s b_s X_s (1-s_s) r_s P_n K_s\} + P_n I_s \qquad ....(13)$$

Note that since all saving in each region is invested, the absence of inter-regional capital flows implies that all saving in a region is invested within that region, or that

$$I_i = s_i\, r_i\, K_i \qquad ....(14)$$

for each region i.

In the short run levels of capital stock in each region, K_i, are given, and commodity markets clear as a result of price variations. Price changes can be examined by considering movements in $p(= P_s / P_n)$. the terms of trade of the South in terms of the Northern good. Since equation (8) implies

$$\pi = V_s / p^{1-m} \qquad ....(15)$$

where $\pi = W_s / P_s$, and since V_s is fixed, we can also examine price changes by examining changes in π, remembering that a rise in p implies a fall in π. Equation (12) implies, substituting from equations (5), (6) and (10), that

$$\pi = \{a_s - ma_s(1-s_s)ca_nk\} / ms_sa_sb_s \qquad ....(16)$$

where $k = K_n / K_s$, which gives the value of π which clears the market for the Southern good. Using equations (5), (6), (9), (10), we see equation (11) is satisfied when (12) is satisfied. Thus market clearing for the Southern good implies also market clearing for the Northern good, so that (16) is the condition for short-run equilibrium in the world economy. Further, assuming that p rises (or *p* falls) in response to an excess demand for the Southern good, this short-run equilibrium is seen to be stable. The short-run equilibrium value of π is shown, for any value of k, by the lower quadrant of fig. 2. We confine attention to levels of k which ensure *p* >0, since negative values are not economically relevant.

In the long run the stocks of capital are allowed to change due to saving and investment. Assuming away depreciation,we have

$$g_i = I_i / K_i \qquad ....(17)$$

where g_i denotes the rate of growth of k_i, which implies, using equations (14),

$$g_i = S_i r_i \qquad ....(18)$$

using equations (7), (9) and (15) we can write this equation for n as

$$g_n = s_n\{1 - V_nb_n - V_s^{\frac{1}{1-m}} \pi^{\frac{-1}{1-m}} c\}a_n \qquad ....(19)$$

which shows that an increase in π, by reducing p or the Southern terms of trade, and making intermediate goods relatively cheaper, increases the Northern profit rate, and therefore implies a higher rate of capital accumulation. The relation is shown in the upper quadrant of Fig. 2 by the curve g_n. Similarly, using equations (10) and (15), (19) for the South can be written as

$$g_S = s_S a_S V_S^{\frac{1}{(1-m)}} \left\{ (1 - \pi b_S) / \pi^{\frac{1}{(1-m)}} \right\} \qquad \text{......(20)}$$

which shows that an increase in π, by reducing p or the Southern terms of trade, and thereby increasing the product wage and the relative price of investment goods, reduces the Southern rate of profit and therefore the rate of Southern capital accumulation. This relation is shown in the upper quadrant of Fig. 2 by the curve g_s. This upper part of the diagram can be compared to Kaldor's (1979) diagram of growth in the world economy with the difference that we measure π rather than p on the horizontal axis.

The movement of the world economy over time can then be examined as follows. Given stocks of capital in the short run imply a given k, say k_1. For k_1, the short-run equilibrium value of π is read off from the lower quadrant, as π_1. Corresponding to π_1 the curves g_n and g_s give the rates of growth of capital in each region. Since at π_1, $g_n > g_s$, and since

$$dk / dt = (g_n - g_S)k$$

k rises over time, with π falling and hence g_n falling and g_s rising over time, until long-run equilibrium at k* is reached. A similar argument holds if we start from a k higher than k*. Given appropriate parameter values we can thus see that the long-run equilibrium, at which the two regions accumulate capital, and hence experience output growth, at the same rate, will exist, and will be stable.

Let us now introduce technological change into the analysis. For this purpose we assume that the search for raw-material substitutes will be more intensive when the relative price of such raw materials increases. Moreover, we assume that technological change can be represented by reductions in c offset by reductions in a_n, that new technology will be adopted by firms if and only if at given prices (since firms are competitive), the rate of profit, r_n, rises,[8] and that such changes are irreversible. Suppose, now, that we start with a level of k, say k_1, which is less than the initial long-run equilibrium level given by k*. Over time, k must rise, so that π must fall, or p must rise. This rise in p will induce firms to search for new technology, as discussed above, and they will adopt the new technology with lower c and a_n (we do not discuss the microeconomics of this) if at given p it raises r_n. This has two implications for the curves in Fig. 2. First, the g_n curve is pushed up since r_n rises at the given p and since $g_n = s_n r_n$.[9] Second, as equation (16) shows, the line in the lower quadrant of fig. 3 will rotate anti-clockwise. Because of the technological change, for given stocks of capital, there will be less of the Northern good produced, and for this, and for the material-saving nature of Northern production, the demand for Southern goods will fall, implying a fall in π.

Fig. 2 shows that the short-run implication of this change, that is, for a given k, there is a rise in π, and a corresponding rise in g_n and a fall in g_s.[10] In the long run, k begins to rise from k*, and keeps rising (in the absence of further changes in technology), till the new long-run equlibrium is reached at k**. This rise implies that capital grows faster in the North than in the South, representing a widening of the gap between the two regions or what may be called uneven development (see Dutt, 1990). Such uneven development, however, is not inevitable in this model. If more rapid Northern growth causes increases in the Northern real wage, as discussed by Emmanuel (1972), the g_n curve can in fact shift down, causing the long-run equilibrium k to fall. Regarding the terms of trade, along the path to the new long-run equilibrium, as k increases, π must fall, which implies that the Southern terms of trade, p, must rise. This rise in p may induce the search for new technology (as discussed above) which, if adopted, will lead to a repitition of the process just described. It may be noted that technological change of the type examined here must initially cause a fall in the Southern terms of trade, but over time will lead to its improvement. If one compares two long-run equilibria, one before and one after a technological change, one will observe an *improvement* in the southern terms of trade. This implies that if we start with a k below k*, the initial long-run equilibrium (like k_1), the Southern terms of trade will improve as result of technological change.[11]

The end of uneven development?

Now we turn to the issue of the NICs. There is no question that the NICs have 'emerged'. This is not inconsistent with the North–South models. Individual Southern regions can develop. Some models appear to suggest that this cannot happen. For instance, some of the models in Dutt (1990) suggest that increases in the Southern saving propensity, s_s., will only imply Southern terms of trade deterioration, and no alteration in k, the ratio of Northern to Southern capital stock in the long run. But this leaves room open for some Southern regions to get ahead by saving at a higher rate-without turning the terms of trade against them. It is worth pointing out that the likelihood is greater for smaller LDCs.

Why some countries can get ahead is a central issue of development economics, perhaps the central issue for many LDCs today. To understand this one needs to study carefully the institutional characteristics of successful countries and compare them to those of less successful countries. The macroeconomic nature of the North-South models makes them not terribly useful tools for this purpose. But they can be used to analyze another important question–the implications for the success of NICs for the rest of the South. In particular, does the emergence of the NICs imply that the end of uneven development is close at hand–will all Southern economies join the ranks of the NICs?[12]

We have already seen that poor nations have grown more slowly than richer nations. This has occurred during a period in which the NICs have grown rapidly: South Korea at 7.1, Hong Kong at 6.2, Taiwan at 6.7, and Singapore at 6.5 per cent.[13] It appears, then, that the rapid growth of the NICs has been accompanied by the North growing faster than most of the South. While this is suggestive that uneven development in this sense is linked to NIC growth, it cannot be used to argue that the high growth of the NICs *caused* uneven development between the North and the South, since growth rates are affected by a variety of other influences (including policy changes). A theoretical framework is needed to analyze these questions, and a North-South model extended to incorporate a NIC does just that.[14]

We consider a model of the world economy with three regions: the North, the newly-industrializing countries and the rest of the South (henceforth the South), which we denote by n, m (for middle) and s. Each region is completely specialized in the production of one good (which may be thought of as a composite commodity) with given technology.

In the North firms produce an investment-cum- consumption good with capital goods and labour using a fixed coefficients production function. The industrial structure in the North is oligopolistic, and firms set the product price as a markup on labour costs, so that

$$P_n = (1+z)W_n b_n \tag{21}$$

where P_i is the price of the product, W_i the money wage, b_i the labour-output ratio (all for region i), and z the Northern markup which is given by the Northern degree of monopoly a la Kalecki. Firms do not utilize their capital fully and adjust their output to meet demand rather than change price which, according to (21), is determined by cost and the markup rate. Firms plan to invest more when the rate of profit is higher (reflecting higher profit expectations or easier financial conditions) and when the degree of capacity utilization is higher. We write the investment function in the form

$$I_n/K_n = g_n(r_n, X_n/K_n) \qquad ....(22)$$

where both partials are positive, I_i is investment, K_i is capital, X_i is output, and r_i the rate of profit, all for region i, and the rate of profit for the North is defined as

$$r_n = (P_nX_n - W_nb_nX_n)/P_nK_n \qquad ....(23)$$

There are two classes in the North, capitalists who own the capital and receive profit income (which the firms distribute to them without retaining any portion of it), and workers who provide labour and earn wage income. Capitalists save a fraction s_n of their income. Workers are available in unlimited supply at a given money wage W_n, and consume all their income. Workers and capitalists spend a constant fraction of their consumption expenditure on the Northern, NIC and Southern goods, where the expenditure shares are given as α_n, β_n, and τ_n, where $\alpha_n + \beta_n + \tau_n = 1$.[15]

The NICs and the South have several features in common. In both regions firms produce a consumption good with capital goods and labour, using fixed coefficients production functions. They operate in competitive markets in the sense that the price of their product varies to clear the market with their production at full capacity, which implies that

$$X_i = a_i K_i \qquad ....(24)$$

where a_i is the output-capital ratio for region i. Firms invest all the savings they can lay their hands on, propelled to do so perhaps by government policy. Like the North, both regions have two classes: capitalists and workers. Capitalists save a fraction of their income, s_i, and workers consume all of theirs.

However, because of the higher level of development of the NICs, the structure of the two regions differs in two respects. First, unlimited supplies of labor imply a subsistence wage in the South (in terms of the Southern good, the only good consumed by Southern workers), V_s., which is fixed, in a subsistence sector not considered explicitly in the model. In the NICs, however, although there is still a reserve army of workers, real wages have increased dramatically as a result of growth. In this context, rather than assuming that the real wage is fixed, it is more appropriate to assume that the labor share in the value of output is given, in a Marxian manner, by the state of class struggle between workers and firms. Denoting the given rate of exploitation or the profit share by ϵ, we have

$$(P_m X_m - W_m b_m X_m) / P_m X_m = \epsilon \qquad ....(25)$$

This shows that a rise in labor productivity in the NICs increases the product wage of workers (which is invariant with respect to productivity change in the South). Second, only domestic goods are consumed in the South, while in the NICs workers and capitalists spend on Northern, NIC and Southern goods in the fixed ratios α_m, β_m and τ_m, where $\alpha_m + \beta_m + \tau_m = 1$. In the South the consumption of imported goods is prevented by government restrictions, but in the NICs, improved export performance has allowed the liberalization of imports of consumption goods.

The three regions engage in trade with each other. For simplicity we assume away the international borrowing and lending and the international mobility of capital, and deal with the situation of balanced trade. This is to deal with the problem of North-NICs-South trade in the simplest possible environment, without introducing the complication of international factor movements. We also assume a regime of fixed exchange rates, setting all exchange rates equal to one.

In the short run the stock of capital is given in each region, and the three commodity makets clear: the Northern goods market clears through variations in Northern capacity utilization, while the NIC and Southern goods markets clear through variations in their prices. Market clearing for the Northern, NIC and Southern markets, respectively, implies,

$$P_n X_n = \alpha_n [1+(1-s_n)z] W_n b_n X_n + \alpha_m (1-s_m \in) P_m X_m + P_n (I_n + I_m + I_s), \qquad \text{......(27)}$$

$$P_m X_m = \beta_n [1+(1-s_n)z] W_n b_n X_n + \beta_m (1-s_m \in) P_m X_m \qquad \text{.......(28)}$$

and

$$P_s X_s = (1-\alpha_n-\beta_n)[1+(1-s_n)z] W_n b_n X_n + (1-\alpha_m-\beta_m)(1-s_m \in)] P_m X_m$$
$$+ [V_s b_s + (1-s_s)(1-V_s b_s)] P_s X_s \qquad \text{.......(29)}$$

Balanced trade in the three regions, which in the absence of capital movements always holds, implies

$$(1-\alpha_n)[1+(1-s_n)z] W_n b_n X_n = \alpha_m (1-s_m \in) P_m X_m + P_n (I_m + I_s) \qquad \text{.......(30)}$$

$$P_n I_m + (1-\beta_m)(1-s_m \in) P_m X_m = \beta_n [1+(1-s_n)z] W_n b_n X_n \qquad \text{.......(31)}$$

$$P_n I_s = (1-\alpha_n-\beta_n)[1+(1-s_n)z] W_n b_n] X_n + (1-\alpha_m-\beta_m)[(1-s_m \in)] X_m P_m \qquad \text{.......(32)}$$

of which only two are obviously independent.

Equations (27) through (32) imply that saving equals investment in each region, or

$$s_n [z/(1+z)] X_n / K_n = I_n / K_n \qquad \text{.......(33)}$$

$$s_m \in (P_m / P_n^{`})] a_m = I_m / K_m \qquad \text{.......(34)}$$

$$s_s (1-V_s b_s)(P_s / P_n)] a_s = I_s / K_s \qquad \text{.......(35)}$$

The determination of short-run equilibrium may now be shown as follows. Substituting (21), (22) and (23) into (33) we get

$$s_n[z/(1+z)]u = g_n[zu/(1+z),\ u] \qquad ....(36)$$

Where $u = X_n/K_n$. Substitution of (33), (34) and (35) into (27) implies

$$s_s(1-V_s b_s)\, a_s\, k_s\, p_s + [s_m \in + \alpha_m(1-s_m \in)]a_m\, k_m\, p_m = (1-\alpha_n)[1-s_n z/(1+z)]u \qquad ....(37)$$

Where $k_i = K_i/K_n$ and $p_i = P_i/P_n$, for i=m,s. Finally, (28) implies

$$[1-\beta_m(1-s_m \in)]a_m\, k_m\, p_m = \beta_n[1-s_n z/(1+z)]u \qquad ....(38)$$

Equation (36) solves for the short-run equilibrium value of u, u*.[16] Once u* is thus solved for, its value can be substituted into (38) to solve for

$$p_m^* = \Sigma_m/k_m \qquad ....(39)$$

where

$$\Sigma_m = \beta_m[1-s_m z/(1+z)]u^*/[1-\beta_m(1-s_m e)]a_m$$

Finally, these values can be substituted into (37) to find

$$p_s^* = \Sigma_s/k_s \qquad ....(40)$$

where

$$\Sigma_s = \frac{\{(1-\alpha_n)\,[1-\beta_m(1-s_m \in)]-[s_m \in + \alpha_m(1-s_m \in)]\beta_n\}\,[1-S_n z/(1+z)]u^*}{s_s\,[1-\beta_m(1-s_m \in)]\,(1-V_s b_s)\, a_s}$$

Since α_i,$\in$ and s_i are all fractions, economically meaningful short-run equilibrium values of p_m^* and p_s^* will be obtained. It can be shown, assuming that u responds positively to excess demand for N-good, and p_i responds positively to excess demand for the i-good for i = m and s, that when the conditions for the existence of short-run equilibrium are satisfied, the equilibrium will also be stable.

In the long run we assume that the stocks of capital, K_i change according to the level of investment in each region, and assume away complications due to depreciation. As mentioned above, technology is assumed to be given in each region. We also assume that the short-run equilibrium conditions are always satisfied when the economy moves over the long run.

The long-run dynamics of the world economy can be examined by analyzing the movement over time of the two short-run state variables, k_m and k_n. By definition.

$$\hat{k}_i = \hat{K}_i - \hat{K}_n \qquad (42)$$

for i = m and s, where overhats denote time-rates of growth. Since $K = I_i/K_i$, this implies using (33) through (35) and (39) and (40), that

$$\hat{k}_m = s_m \in a_m \Sigma_m / k_m - s_n [z/(1+z)] u^* \qquad (43)$$

and

$$\hat{k}_s = s_s (1 - V_s b_s) a_s \Sigma_s / k_s - s_n [z/(1+z)] u^* \qquad (44)$$

where Σ_1 have been defined above.

Equations (43) and (44) show that over the long run k_1 (i=m,s) moves over time depending only on its own level. Since k falls with k_i it follows that k_i will converge to a position at which k_1=0 and we may call this the long-run equilibrium position of the would economy. At it, we see from (39), (40), (43) and (44) that

$$k_m^* = \frac{[\{1+(1-s_n)z\} s_m \in \beta_n]}{\{1-\beta_m(1-s_m \in)\}\, s_n z} \qquad \text{.......}(45)$$

and

$$k_s^* = \frac{[(1-\alpha_n)\ \{1-\beta_m(1-s_m \in)\} - \{\alpha_m + s_m \in (1-\alpha_m)\}\beta_n]\ \{1+(1-s_n)z\}}{\{1-\beta_m(1-s_m \in)\}\, s_n z} \qquad \text{.......}(46)$$

where the starred levels denote long-run equilibrium levels. From the definitions of k_m and k_s, it also follows that

$$K_m^* / K_s^* = s_m \in \beta_n / \{(1-\alpha_n)[1-\beta_m(1-s_m \in)] - [\alpha_m + s_m \in (1-\alpha_m)]\beta_n\} \qquad \text{.......}(47)$$

These levels are the long-run equilibrium levels of ratios of capital stocks of the three regions. If parametic shift occurs to increase one of these ratios we may conclude that the region for which the capital stock is given in the numerator of the ratio has grown faster, as a result of the shift, than the region for which the capital stock is given in the denominator.

We now examine the consequences of parametric shifts to see what effect changes in parameters which plausibly explain NIC growth have on the North and the rest of the South.

More specifically, we are concerned with shifts which cause increases in the long-run equilibrium values of k_m and K_m/K_s and their effects on evenness of North-South development as measured by k_s. For this purpose we consider increases in the NIC savings (s_m) and exploitation $\in$ rates, and in the shares of Northern and NIC spending on NIC goods (β_i), all of which have been argued to be important causes for the rapid growth of the NICs. The short-run effects of these parametric shifts on u, p_m and p_s can be found from equations (36), (39) and (40), and the long-run effects can be found from equations (45) through (47).

An increase in the NIC saving rate, s_m, leaves the Northern rate of growth and capacity utilization rate unaffected. In the short run it reduces the equilibrium value of p_m by shifting NIC demand away from the NIC good. The fall in NIC demand for the Southern good as a result of the terms of trade loss and the lower level of consumption spending also reduces p_s. The rise in the NIC saving rate, despite the fall in the NIC terms of trade, increases the rate of NIC rate of accumulation, but the fall in the Southern terms of trade implies a reduction in its rate of accumulation. Consequently, in the long run there is a rise in k_m and a fall in k_s, as confirmed by the derivatives

$$dk_m^*/ds_m = [1+(1-s_n)z] \in \beta_n(1-\beta_m)/s_n\, z[1-\alpha_m(1-s_m\in)]^2$$

$$dk_s^*/ds_m = -[1+(1-s_n)z] \in \beta_n\tau_n/s_n\, z[1-\beta_m(1-s_m\in)]^2$$

It immediately follows that an increase in s_m also reduces the long-run equilibrium level of K_m/K_s. The higher savings rate makes the NICs save and invest and therefore grow faster than the rest of the world. The higher saving rate also shifts the world composition of demand towards the investment good, which is the Northern good, and thus makes the North grow relatively faster than the South.[17]

A change in the NIC rate of exploitation has the same effects as a change in the NIC savings rate, since what the rise in this parameter does is redistribute income from workers to capitalists, thereby increasing the rate of savings[18]. Thus the NICs grow most rapidly, and there is North-South uneven development.

Turning next to changes in expenditure shares, the growth of the NICs can be explained by shifts in the patterns of spending favouring the NICs, that is, by increases in β_n and β_m. Such inceases can come about as a result of reductions in α_i (at the expenses of Northern goods) or τ_i (at the expense of Southern goods) or both.

A rise in β_n can be taken to occur as a result of successful export-promotion policy by the NICs. (The question as to whether free markets or active state brought about such changes need not detain us). If it is accompanied by a compensating fall in τ_n which leaves α_n unchanged, it shifts demand from the Southern good to the NIC good. In the short run there is an improvement in the NIC terms of trade, and a deterioration of the Southern terms of trade; these are translated into a higher growth rate for the NICs and a lower growth rate for the South. Since the Northern growth rate, which depends only on its saving and investment parameters, remains unchanged, the long run effect is an increase in NIC capital stock relative to the North's and a fall in the Southern capital stock relative to the North's. This is confirmed by

$$dk_m^*/d\beta_n = [1+(1-s_n)z]s_m \in /s_n z[1-\beta_m(1-s_m)\in]$$

$$dk_s^*/d\beta_n = -[\alpha_m+s_m\in(1-\alpha_m)]\,[1+z(1-s_n)]/[1-\beta_m(1-s_m\in)]s_n z$$

There is clearly a rise in the long-run equlibrium value of K_m/K_s. If, on the other hand, the increase in β_n is accompanied by a fall in α_n, leaving τ_n unchanged, so that there is a shift in demand from the Northern to the NIC, there is a shor-run, improvement in the NIC terms of trade, and also an improvement in the Southern terms of trade due to greater demand from the NICs. Thus both NIC and Southern growth rates increase in the short run, so that both k_m and k_s increase in the long run. The change in k_m^* is given by the expression for the previous case, while the change in k_s^* is given by

$$dk_s^*/d\beta_n = \tau_m(1-s_m\in)[1+z(1-s_n)\}/\{1-\beta_m(1-s_m\in)\}s_n z$$

This will be negligible if the NIC imports very little from the South. Though both k_m and k_s increase in the long run, K_m/K_n increases in this case as long as $\tau_n > 0$. In both cases, therefore, the NICs will growth faster than the North and South; what happens to the South relative to the North depends on whether the increase in NIC exports to the North is primarily the result of a reduction in Northern spending on Northern or Southern goods;k_s is more likely to fall in the long run if the shift towards the NIC goods is mainly at the expense of the Southern good If t is small (the NIC imports from the South are negligible), and there is *some* reduction of τ_n when β_n rises, k_s must fall in the long run.

A rise in β_m can represent successful import substitution in the NICs. If the increase in β_m is accompanied by a fall in τ_n but no change in α_m, along the lines discussed in the previous paragraph there is a short-run increase in g_m and decrease in g_s. Since g_n is unchanged, the long-run equilibrium value of k_m rises and k_s falls, as confirmed by the expressions

$dk_m^* / d\beta_m = [1+(1-s_n)z]s_m\in \beta_n(1-s_m\in)/s_n z[1-\beta_m(1-s_m\in)]^2 > 0$
$dk_s^* / d\beta_m = [1+(1-s_n)z]\beta_n[\alpha_m+s_m\in(1-\alpha_n)]/s_n z[1-\alpha_m(1-s_m\in)]^2 < 0$ and the ratio K_m/K_s also increases. If, on the other hand, the increase in β_m is accompanied by a fall in α_m and no change in τ_m, in the short run both NIC and Southern growth rates would increase, so that in the long run k will increase as shown in the above expression for $dk_m^* / d\beta_m$, and k_s^* will rise as shown by

$$dk_s^* / d\beta_m = \tau_m \beta_n(1-s_m\in)^2[1+z(1-s_n)]\beta_n/s_n z[1-\beta_m(1-s_m\in)]^2 > 0.$$

K_m/K_s will rise with a rise in β_m in this case if $\tau_n>0$. Our conclusions here are similar to that obtained for the case of a rise in β_n.

Two factors determine whether a rise in β_i will increase or reduce k_s^*. First, it depends on whether the increase comes about primarily as a result of a reduction in α_i or that in τ_i. One view is that the NICs compete mainly in the production of industrial goods, in which they have a comparative advantage, so that they are mainly in competition with the North, and not the rest of the South which has a comparative advantage in the production of an export of primary goods. This implies that increases in β_i are primarily reflected in decreases in α_i. While there is certainly some truth in this claim, the idea that non-NIC less developed economies are specialized in the production of primary goods is an exaggeration, with these economies increasingly exporting manufactured products.[19] The main argument in favor of the alternative view, that increases in β_l are mainly reflected in reductions in τ_l, is that the growth of exports is not so much explained by comparative advantage, but by marketing, product standardization, and quality, and countries which have greater experience in exporting are more difficult to compete against. Thus, NIC export growth is likely to relatively crowd out, especially in the later stages, other Southern exporters of manufactured goods. Which of these two effects is stronger can only be ascertained after a detailed empirical investigation of the question. Second, it depends on the size of τ_m; with a negligible τ_m if a rise in β_l is accompanied by even some reduction in α, there will be a fall in k_s^* (because with a small τ_m we have found that the increase in k_s^* due to a fall in α_l is small). In fact the NICs have relatively small imports from the rest of the South. Overall, therefore, we may expect that an increase in β_l is accompanied by a fall in k_s^*.

Our conclusion is that it is very likely that changes which have made the NIC grow faster most likely have led to a widening gap between the North and the rest of the South. If the ability of the latter to grow depends on their relative level of development, this means that NIC growth makes it less likely for them to develop, and the longer they wait, the more difficult. Of course, our model must be supplemented by using the commodity composition of exports of the three regions, and changes in it, which are not available at the level of aggregation required for our analysis.

Conclusion

The objective of this paper was to provide a flavor of some North-South models and of the issues relating to uneven development they can address. It is hoped that the paper shows that the models are useful, but that they must be supplemented by a detailed examination of history, institutions, and the structural differences between rich and poor regions, and data on trade composition for the North, South and NICs. For the question of the origins of uneven development, models have to be developed in the light of the economic history of relations between different regions. For the question of the perpetuation of uneven development, one must study carefully the structures of rich and poor regions and the causal mechanisms of possible uneven development to which they give rise before embarking on model construction. Finally, for the question of NIC emergence, the detailed study of institutional characteristics of particular economies and data on trade flows need to supplement North-South models.

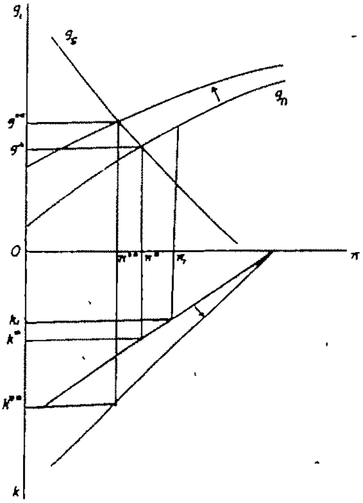


Fig.-2

Footnotes

1. The data is taken from the *World Development Report*, 1992, and refer to GNP growth rate between 1965 and 1990. The high growth rate for the low-income countries as a whole is explained by the rapid growth of China (at 5.8 per cent) which, as a socialist economy undergoing radical policy changes, is a special case

2. It should be pointed out that uneven development in the observed sense may not be related to uneven development in this theoretical sense, because even if uneven development is observed, interactions between rich and poor nations may in fact be to some extent moderating, and because if uneven development is not observed, interactions between the nations cause uneven development which is reversed by developments within the nations

3 For a formal analysis of this and other related models see Dutt (1986), which also allows agricultural productivity and the output-capital ratio in manufacturing, to rise. Changes in agricultural productivity can stem the forces of uneven development, but if the parameter like τ for agriculture (representing spillover effects from the manufacturing sector) is smaller than τ for manufacturing, uneven development still occurs.

4. This discussion draws on Dutt (1992b).

5 This mechanism has been relatively neglected in the literature, with most models assuming that both regions export final, rather than intermediate goods. Exceptions include Kemp and Ohyama (1978) which examines the static implications of such a trade pattern and Mainwaring (1989, 1991) which discusses cyclical, rather than secular, tendencies. The following discussion follows Dutt (1993)

6 The implicit assumption is that, perhaps because of Northern colonialism in the South, Southern producers are unable to compete in international markets as producers of consumption goods. The Southern good can thus be consumed inside the South, where it is protected by (say) transport costs which are not explicitly modeled in the paper. The Southern good is thus thought of as a composite good The Southern raw material can be used as a low-quality consumption good (in the South) after very little processing (which is conveniently ignored), or exported and then processed by Northern industry

7. We assume that money magnitudes are measured in terms of a common currency, or that the exchange rate is fixed at unity.

8. An alternative may be to assume that Northern producers consider expected future prices in deciding on whether to adopt the new technology. It can be argued that if they expect material prices to rise in the future, they may adopt the new technology even if the rate of profit, r_n, in terms of current prices, falls This will imply a downward shift in the g_n curve, unlike what is discussed below. However, this will increase the equilibrium π, and thus reduce the equilibrium p, contradicting the expectation of an increase in p. For this reason we choose to assume that Northern producers will adopt the new technology only if it is more profitable at existing prices.

9 This statement is strictly true only at the prevailing p, and hence at the prevailing π. However this does not creat a problem for our subsequent analysis, as can be seen as follows. Equations (9) and (14) for n imply that $g_n = s_n (1-V_n b_n-cp)a_n$

so that, totally differentiating with respect to changes in c and a_n at a given p, we get

$$dg_n = s_n\{(1-V_nb_n-cp)da_n-pa_ndc\}$$

The kind of technological change we are concerned with implies dc< 0 and $da_n< 0$ At the existing p, since $dr_n> 0$, $dg_n>0$, the g_n curve must be higher at it due to the technological change. At p higher than the existing p (which implies that at π lower than the existing one), the second term (which involes dc) becomes larger and the first term (involving da_n) becomes smaller, to that $dg_n> 0$ will still hold At p lower than the existing p (or π higher than existingπ) the opposite happens, so that it is possible (though not certain), that $dg_n< 0$. Since, as we shall see below, technological change will involve a movement of the economy to the left of the current position, or p must be smaller at the new long-run equilibrium position, our results will carry through

10 The rise in g_n is not a certainty after π increases, since it is possible, as noted above, that the g_n curve may be lower at high levels of π after the technological change than before it. Since π is higher than before, and this causes g_n to rise, other things constant, it is most likely that g_n will rise in the short run.

11. For an analysis of what happens when capital is internationally mobile, see Dutt (1993).

12. A model developed by Mainwaring (1990) seems to imply just that

13. The data is taken from *World Development Report* 1992 for all countries except for Taiwan, for which the rate has been calculated from *Statistical Yearbooks of the Republic of China*. The data refer to the period 1965-90.

14. This follows Dutt (1992a).

15. The structure of the North basically follows Taylor's (1981, 1983) formulation In Mainwaring's (1990, 1991) analysis for the relevant stage, Northern growth is constrained by labor supply growth rather than by effective demand as assumed in Taylor's framework.

16. Sufficient conditions for an economically meaningful solution to exist are $g_n(0,0) > 0$ (which is satisfied if there is a positive, autonomous part to Northern investment), and $s_nz/(1+z) > g_{n1}z/(1+z) + g_{n2}$, where g_{n1} is the it derivative of the g_n function (which requires that saving is more responsive to changes in capacity utilization than is investment, and is a standard

stability condition for all Keynesian quantity adjustments models) We assume that these conditions are satisfied and also that the saving and investment functions in the North, and z, are such as to yield u* less than the maximum possible degree of capacity utilization possible with the given technology

17 It may be noted that we do not require the assumption that the North produces and exports only Investment goods to obtain this result.

18. Since in our model NIC workers and capitalists do not have different consumption spending patterns for different goods, effects which can arise from these differences are being ignored

19 According to the *World Development Report* for 1990, in 1988, 53 per cent of the exports of low-income economies consisted of manufactures Leaving out China and India from this group, the percentage falls considerably to 25, but it is still clearly substantial.

REFERENCES

Arthur, Brian (1990). "Positive Feedbacks in the Economy", *Scientific American*, February, 92-99

Basu, B.D. (1935). *Ruin of Indian Trade and Industries*, 3rd edition, Calcutta: R Chatterjee.

Bienefeld, M. (1981). ""Dependency and the Newly Industrializing Countries (NICs): Towards a Reappraisal", in D. Seers, ed , *Dependency Theory, A Critical Assessment*, London: Pinter

Cairncross, Alex K. (1960) " International trade and economic development", Kyklos, 13(4), 545–58.

Chandra, Bipan (1966). *The Rise and Growth of Economic Nationalism in India*, New Delhi· People's Publishing House

Charlesworth, Neil (1982) *British Rule and the Indian Economy*, 1800-1914, London. Macmillan

Chaudhuri, K.N. (1982) ""Foreign Trade and the Balance of Payments (1757-1947)" in D.Kumar, ed., *The Cambridge Economic History of India*, Vol. 2, Cambridge University Press.

Darity, W A , Jr. (1982). " A General equilibrium Model of the Eighteenth-Century Atlantic Trade", Research in Economic History, 7, pp. 287-326, Greenwhich, Conn · JAI Press.

Dutt, Amitava K (1986). " Vertical Trading and Uneven Development", *Journal of Development Economics,* (20), 339-59.

Dutt, Amitava K (1988a). "North-South models a critical survey", unpublished, University of Notre Dame.

Dutt, Amitava K (1988b) "Inelastic demand for Southern goods, international demonstration effects and uneven development", *Journal of Development Economics*, 29 (1), July, 111-22

Dutt, Amitava K. (1990) *Growth, distribution and uneven development*, Cambridge: Cambridge University Press

Dutt, Amitava K (1992a). "The NICS, global accumulation and uneven development", *World Development*, 20(8), August, 1159-71.

Dutt, Amitava K. (1992b) "The origins of uneven development The Indian Subcontinent", *American Economic Review*, 82 (2), May, 146-50

Dutt, Amitava K (1993). "Southern primary exports, technological change and uneven development", forthcoming, *Cambridge Journal of Economics.*

Dutt, Romesh C. (1901, 1903). *The Economic History of India*, Vols. I and II, London: Routledge & Kegan paul, 1901, 1903.

Findlay, Ronald (1980). "The terms of trade and equilibrium growth in the world economy", *American Economic Review*, 70 (3), June 291-9.

Findlay, Ronald (1992). The Roots of Divergence: Western Economic History In Comparative perspective, *American Economic Review*, 82 (2), May, 158-161.

Kaldor, Nicholas (1979). "Equilibrium theory and growth theory", in M.J.Boskin, ed , *Economics and Human Welfare*, London and New york: Academic Press.

Kemp, Murray and Ohyama, Michihiro (1978). "On the sharing of trade gains by resource-poor and resource-rich countries", *Journal of International Economics*, 8,93-115

Krugman, Paul (1981) "Trade, accumulation, and uneven evelopment", *Journal of Development Economics*, (8), 149-161

Krugman, Paul (1991). *Geography and Trade*, Cambridge, Mass. · MIT Press.

Maddison, Angus (1971) *Class Structure and Economic Growth*, New York. W W. Norton

Mainwaring, Lynn (1989). "World income shares and the terms of trade A North-South growth cycle", *Bulletin of Economic Research*, 41(3), 185-96.

Mainwaring, Lynn (1990). "Endogenous NIC-formation in a North-South Growth Model", Mineo, Swansea, UK: University College of Swansea.

Mainwaring, Lynn (1991) *Dynamics of uneven development*, Aldershot, Hants . Edward Elgar

Molana, Hasan and Vines, David (1989) "North-South growth and the terms of trade. A model on Kaldorian lines", *Economic Journal*, 99

Nurkse, Ragnar (1959). *Patterns of Trade and Development*, Stockholm· Almquist and Wicksell. Republished, 1961, Oxford. Blackwell.

Raychaudhuri, Tapan (1968) "A Re-interpretation of Nineteenth Century Indian Economic History?" *Indian Economic and Social History Review*; (5),77-100.

Raychaudhuri, Tapan (1982). "Non-Agricultural Production – Moghul India", in T. Raychaudhuri and I Habib, eds , *The Cambridge Economic History of India*, Vol. 1, Cambridge: Cambridge University Press

Romer, Paul (1986). "Increasing returns and long-run growth", *Journal of Political Economy*, 94 (5), October, 1002-37.

Rostow, Walt W. (1980) *Why the Poor Get Richer and the Rich Slow Down*, Austin. University of Texas Press.

Sabel, Charles and Zeitlin, Jonathan (1985), "Historical Alternatives to Mass Production· Politics, Markets and Technology in Nineteenth-Century Industrialization", *Past and Present*, 133-76.

Singer, Hans (1987). "Terms of trade and economic development", in J Eatwell, M Milgate and P Newman, eds., *The New Palgrave*, London· Macillan.

Taylor, Lance (1981) "South-North trade and Southern growth· bleak prospects from a structuralist point of view", *Journal of International Economics*, 11(4), November, 589-602.

Taylor, Lance (1983). *Structuralist Macroeconomics*, New York: Basic Books

Vines, David (1984). "A North-South growth model along Kaldorian lines", mimeo, London: Centre for Economic Policy Research.

Wilson, Horace Hyman (1845) *History of British India. From 1805 to 1835*, Vol. I, London· James Madden and Co.

Quantitative Economics

3

Granger Causality and Misspecified Vector Autoregressions*

Kajal Lahiri & Nlandu Mamingi

Introduction

According to Granger (1969), a variable x_t does not cause another variable y_t if the past of x_t does not help better predict y_t than does the past of y_t alone. This null hypothesis and others of the same type have been tested using different formulations [e.g., Granger (1969), Sims (1972), Sargent (1976), and Geweke et al. (1983)]. In the Granger-Sargent formulation the null hypothesis is generally tested with an *F* statistic in the vector autoregression (VAR) framework, i.e.,

$$A(B)\, Z_t = \varepsilon_t \tag{1}$$

where *B* is the backward shift operator, $Z_t = (x_t, y_t)'$ is a vector of variables and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ represents two uncorrelated white noise errors.

It is important to point out that covariance stationarity of variables is required to implement equation (1). Yet in the real world economic variables behave more like integrated processes than stationary ones [see Nelson and Plosser (1982)]. Thus, researchers typically first difference the data before implementing the test. With first diffeenced variables, model(1) becomes:

$$A(B)\,(1-B)\, Z_t = \varepsilon_t \tag{2}$$

It has been found that if x_t and y_t are individually integrated of order one but cointegrated, that is, their linear combination

$$u_t = \delta' Z_t \tag{3}$$

---

* An earlier version of this paper entitled "Causality Testing in Vector Autoregressions and Error Correction Models" was prepared for presentation at the 1992 Indian and South-East Asia meeting of the Econometric Society in Bombay. The views expressed here are our own, and do not necessarily represent those of the World Bank, the Social security Administration and the U.S Department of Health and Human Services.

is stationary with cointegrating vector δ', then Granger causality is adequately detected not in the vector autoregression model (2) but in an error correction model (ECM):

$$A^*(B)\,(1-B)\,Z_t = -\gamma\delta' Z_{t-1} + \varepsilon_t \qquad (4)$$

where $A^*(0) = A(0) = I_N$ and $\delta' Z_{t-1}$ is the lagged error correction term. The reason is that the relationship (2) is misspecified with respect to (4): the lagged error correction term that captures the long run equilibrium is simply missing in the VAR formulation (2). [see Engle and Granger (1987, p.259)].

The problem of misspecification in the first differenced VAR has not yet received sufficient attention in the literature although there are many empirical studies which faced this problem[1]. Wickens and Breusch (1988), Sims, Stock and Watson (1990), Spanos (1990), Masconi and Gianini (1992) and Toda and Phillips (1993) and others have recently helped to pin point the problem. This is an attempt to fill the gap.

The essay deals with the following precise questions. First, is there any Granger causality distortion due to an omission of an error correction term in an unrestricted VAR in differences estimated with cointegrated variables? Second, to what extent is such a distortion of great concern in real life situations? To put it differently, the likelihood of change in the pattern of Granger causality between variables in a misspecified unrestricted VAR as well as the sensitivity of Granger causality distortion to the degree of "cointegratedness" of variables and to the sample size are the issues raised in this essay.

Before proceeding any further, three remarks are in order. First, the comparison between the two types of causality structures, that is, causality in the ECM and the unrestricted VAR, has to be carried out with caution. Indeed, while causality in the ECM is defined in both short and long terms, that in VAR in differences reflects only a short-run effect. Second, long-run Granger causality in the ECM means changes in the variables caused by the equilibrium error. The reasons for the equilibrium error are not brought in explicitly In Campbell and Shiller (1988), for instance, equilibrium errors are caused by agent's forecast errors. Third, in large samples, the estimates of the parameters of A(B) in equation (2) are asymptotically the same as those of $A^*(B)$ in equation (4) [see Phillips and Durlauf (1986)].

---

[1] See, for example, Anderson et al (1986), MacDonald and Kearny (1987), Christiano and Lyungqvist (1987), Darrat (1987), Atkins (1989), Miller and Russek (1991) and Garcia and Zapata (1991).

The remainder of the paper is divided as follows. Section 2 succinctly introduces the concepts of cointegration, error correction model and Granger causality as well as their relationships Section 3 sets up the experimental design. Section 4 gives the results of the Monte Carlo simulations. Section 5 contains concluding remarks.

Some Concepts

The concept of cointegration was first introduced in economics by Granger (1981) and later developed by Granger and Weiss (1983) as well as Engle and Granger (1987). Let Z_t be an n (n=m+1) × 1 vector of non deterministic variables x_t and y_t where y_t is scalar variate and x_t is a set of m ($m \geq 1$) variables. Z_t is integrated of order *d* at frequency $\Theta = 0, Z_t \sim I(d)$, if each of its components is integrated of order d, that is, has a stationary and invertible ARMA representation after differencing d times. In the frequency domain, $Z_t \sim I_0(d)$ if its spectrum if $f(w) = c(w)^{-2d}$.

The components of Z_t are said to be cointegrated of order (d,b), that is, $Z_t \sim (d,b)$ if all of these components are individually integrated of order d and there exist some linear combinations (i.e., r < n),

$$u_t = \delta' Z_t$$

such that

$$u_t \sim I(d-b)$$

with b > 0. In the bivariate framework with $d = b = 1$, the r linear combinations reduce to one and $u_t \sim I(0)$.

This concept of cointegration has been extended to include other interesting cases. For example, one can have deterministic variables. In such a case, cointegration requires that both deterministic (i.e., trend) and random components not to drift too far apart. Cointegration can well be defined at a frequency other than zero (i.e., seasonal frequencies). Precisely, if x_t and y_t are integrated at both zero and some seasonal frequencies, then cointegration would mean a linear combination that reduces the order of integration of the two variables at a particular frequency. Of course, this rules out the case where the two variables do not have unit roots at corresponding frequencies. [see Hylleberg et al. (1990)]. Another generalization comes from the autoregressive conditional heteroscedasticity (ARCH) literature. Here cointegration means copersistence in the variance.

To sum up, cointegration refers to the long-run equilibrium between variables. In this scenario u_t is simply the extent to which the system is out of equilibrium. In fact, the cointegrated equation provides a framework for modelling both the long-run equilibrium, $u_t = \delta' Z_t$, information about which is often given by economic theory, and the short-run dynamics for which economic theory tells relatively little [Hylleberg and Mizon (1989), p. 117)].

Long-run and short-run elements are explicitly found in the so-called error correction models [see Phillips and Loretan (1991) and the references therein]. Thus, an error correction model (ECM) is either a dynamic model or a VAR model with the characteristics that it takes the previous equilibrium error into account in the explanation of changes in variables. In other words, an ECM is a model that explains a change in a variable by its own past change, the present and/or the past change of another variable (other variables) and a lagged error correction term.

It should be pointed out that the history of the ECM dates back to the early work by Phillips (1954). In fact, because of the difficulty in the estimation of continuous models that he formulated, the ECM was not popular among economists until Sargan (1964) formulated a discrete version in his seminal study on wages and prices in the UK. It is a variant of the latter that Hendry and his associates [Hendry (1976), Davidson et al. (1978) and Hendry and Richard (1982)] used to develop an ECM which is derived following the so-called Hendry methodology. Basically, in this framework, the ECM is a dynamic model which incorporates a lagged equilibrium error term. Another innovation regarding the ECM came from Granger (1981) who used the notion of cointegration and the Granger representation theorem to derive an ECM which is a VAR with a past error correction included in the model. This is the type of ECM used in this essay.

The link between the concepts of cointegration and ECM, has been developed by Granger (1983) in a theorem called "The Granger Representation Theorem" that states that every cointegrated series has a valid error correction model and vice versa. For example, the following ECMs (à la Granger) can be derived from a pair of cointegrated variables

$$\Delta x_t = -\gamma_1 u_{t-1} + \sum_{i=1}^{p} b_i \Delta y_{t-i} + \sum_{i=1}^{p} c_i \Delta x_{t-i} + \varepsilon_{1t} \qquad (5)$$

$$\Delta y_t = -\gamma_2 u_{t-1} + \sum_{i=1}^{p} b_i^{*} \Delta y_{t-i} + \sum_{i=1}^{p} c_i^{*} \Delta x_{t-i} + \varepsilon_{2t} \qquad (6)$$

with at least one $\gamma_i \neq 0$. Note that the corresponding VAR is the model (5)-(6) itself without the lagged error correcting term.

From the literature, it is known that, given relevant information, a variable y_t Granger-causes another variable x_t if the latter variable can be better explained by adding the past values of y_t in its information set. In a VAR formulation, this is usually tested by an F or a Wald statistic in equations (5) and (6) without an error correction term. In the ECM framework, this is done by looking at the significance of the γ_i. In fact, by the Granger Representation Theorem it is known that if two variables are cointegrated then there must exist Granger causality in at least one direction.

Experimental Design

The experimental design aims at answering the following questions. First, is there any Granger causality distortion due to an omission of an error correction in the context of a VAR in first differences estimated with cointegrated variables? Second, is there any relationship between Granger causality distortion and the degree of cointegration of variables? Third, does the distortion depend on the sample size?

To examine these questions, we use the following modified version of the data generation process (DGP) due to Banerjee et al. (1986)

$$\Delta y_t = \alpha(y_{t-1} - x_{t-1}) + \beta_1 \Delta x_{t-1} + \beta_2 y_{t-1} + e_{1t} \tag{7}$$

$$\Delta x_t = e_{2t} \tag{8}$$

where x_t and y_t are $I(1)$ variables, Δ is the first difference operator, e_{1t} and e_{2t} are the errors that are mutually and serially uncorrelated, α is somehow the long-run response, β_1 and β_2 are the short-run parameters and $(y_{t-1} - x_{t-1})$ is the lagged error correction term.

The lagged error correction term depends on the degree of cointegration in the following way:

$$y_{t-1} - x_{t-1} = u_{t-1} = \rho u_{t-2} + e_{3t-1}$$

where the coefficient of autocorrelation, $\rho < 1$, expresses the degree of cointegration and e_{3t} is a well-bahaved error term. As can be seen, as $\rho \to 0$ the variables become highly cointegrated.

To study the sensitivity of our results with respect to the parameters of interest, we choose two sets of parameter values: A_1: $\alpha = -0.7$, $\beta_1=0.5$, $\beta_2=0.6$, $\rho=0.0$, 0.3, 0.6, 0.9 and A_2: $\alpha = -0.5$, $\beta_1 = 0.7$, $\beta_2 = 0.1$, $\rho = 0.0$, 0.3, 0.6, 0.9. These values have been chosen because their corresponding ECMs fit well the DGP. The two sets of parameter values imply that the two variables of interest are cointegrated and there is a unidirectional causality from x_t to y_t. After generating the variables of interest, we fit the following ECM

$$\Delta y_t = c(y_{t-1} - x_{t-1}) + c_1 \Delta x_{t-1} + c_2 y_{t-1} + v_{1t} \tag{9}$$

and VAR

$$\Delta y_t = d_1 \Delta x_{t-1} + d_2 y_{t-1} + v_{2t} \tag{10}$$

$$\Delta x_t = d_3 \Delta x_{t-1} + d_4 y_{t-1} + v_{3t} \tag{11}$$

where the v's are some well-behaved error terms and the other variables have been defined before.

Each experiment is repeated 1000 times. Concretely, first we generate the variables of interest from the DGP. Second, we fit the ECM (9) and the VAR (10)-(11) and compute the statistics of interest for Granger causality comparison. Third, we compute the frequency of reversals of x_t Granger-causing y_t in the ECM into other causal relations in the VAR. The sample size takes values N=50, 100, 200, 400, 1000 and 5000. All the computations are done with the Gauss Programming Language version 3.1.

Note that, in our view, at least three different interpretations of causality can be derived from the ECM (9). First, we can simply compute the t statistic of c and accept the existence of a long-run causality from x_t to y_t if c is significantly different from zero. Second, given the existence of a long-run equilibrium, we can examine a short-run causality from x_t to y_t by checking whether c_1 is significantly different from zero using an *F* statistic. Third, we can test jointly for the existence of long-run and short-run causality, $c \neq 0$ and $c_1 \neq 0$, with an *F* statistic.

Then long-run causality found in the ECM is not present in the VAR formulation (10)-(11). Here the lack of causality from x_t to y_t that is the null hypothesis $d_1 = 0$, is tested with an *F* statistic. The non-causality from y_t to x_t, that is the null hypothesis $d_2 = 0$, is also tested by an F statistic. The F statistic for the ECM and the VAR has the following structure: $F = (RSS_r - RSS_u) \times (n-k) / (RSSu \times r)$ where RSS_r and RSS_u stand for residuals sum of squares in the restricted and unrestricted regressions respectively, n–k is the degrees of freedom for the unrestricted regression and *r* is the number of constraints.

In this exercise, joint long-run and short-run causality in the ECM is compared with short-run causality in the VAR. Since the ECM is supposed to be correctly specified, Granger causality distortion occurs if the unidirectional causality from x_t to y_t found in the ECM is not uncovered in the VAR model. Thus, overall Granger causality distortion is captured by the percentage of Granger causality reversals.

Monte Carlo Results

The following observations result from Table 1. First, it is remarkable that the estimated ECM fits the DGP well to the extent that the correct Granger causality direction is uncovered in most of cases. This information is obtained by adding up rows (1), (2) and (3) for each sample size. The case where Granger causality from x_t to y_t is not uncovered by the ECM is classified in the category called "others". Second, Granger causality distortion (captured by "reversals") does exist and is a serious matter particularly for highly cointegrated variables. Third, it seems that for each sample size, the higher the degree of cointegration, the more likely is the distortion. For example, while with $\rho = 0.9$ the distortion is 44 percent, for $\rho = 0.0$ it reaches 95 percent with a sample size 400. Note that the distinction between highly and weakly cointegrated variables in terms of Granger causality distortion is not too apparent in small samples (see, for example, the results for N = 50). Fourth, the distortion tends to die out as the sample size becomes larger. Thus, with $\rho = 0.9$, the distortion (in percent) is 89, 82, 67,44,10, and 5 for N = 50, 100, 200, 400, 1000 and 5000, respectively. Nevertheless, the most striking result is that the speed of disappearance of distortion largely depends on the degree of cointegration in the sense that the higher the degree of cointegration, the slower the speed of disappearance of distortion.For example, for a sample as big as 5000, while Granger causality distortion is 5 percent for $\rho = 0.9$, the distortion is still more than 90 percent for $\rho = 0.0$. In other words contrary to common belief, Granger causality distortion can be substantial in large samples.

Table 1

Causality Pattern Reversals (in per cent) from ECM to VAR (see equations(9),(10) and (11) with 1000 replications using F statistics at the 5% level of significance:

DGP: $\alpha = -0.7$, $\beta_1 = 0.5$, $\beta_2 = 0.6$, $\rho = 0.0, 0.3, 0.6, 0.9$. → : causality direction

N	Causality from ECM to VAR	$\rho = 0.0$	0.3	0.6	0.9
50	x→y to y→x	4.80	4.70	4.00	3.60
	x→y to feedback	0.30	0.40	0.60	0.40
	x→y to non-caus.	86.20	86.80	88.70	85.30
	x→y to x→y	5.00	5.10	5.70	10.50
	others	3.70	3.00	1.00	0.20
	reversals	94.81	94.74	93.30	89.48
100	x→y to y→x	4.70	4.80	5.00	3.70
	x→y to feedback	0.20	0.20	0.40	0.20
	x→y to non-caus.	90.40	90.10	86.00	77.80
	x→y to x→y	4.70	4.90	8.60	18.30
	others	0.00	0.00	0.00	0.00
	reversals	95.30	95.10	91.40	81.70
200	x→y to y→x	4.70	4.60	4.80	3.90
	x→y to feedback	0.60	0.60	0.60	2.10
	x→y to non-caus.	89.50	88.50	82.30	61.00
	x→y to x→y	5.20	6.10	12.30	33.00
	others	0.00	0.00	0.00	0.00
	reversals	94.80	93.90	87.70	67.00

Contd.

Table : 1 (*Contd.*)

N	Causality from ECM to VAR	$\rho = 0.0$	0.3	0.6	0.9
400	x→y to y→x	5.40	5.50	4.00	1.80
	x→y to feedback	0.20	0.30	1.50	3.10
	x→y to non-caus.	89.90	87.00	73.80	38.90
	x→y to x→y	4.50	7.20	20.70	56.20
	others	0.00	0.00	0.00	0.00
	reversals	95.50	92.80	79.30	43.80
1000	x→y to y→x	4.80	4.90	3.20	0.30
	x→y to feedback	0.30	0.50	2.00	5.10
	x→y to non-caus	90.40	84.60	48.40	5.10
	x→y to x→y	4.50	10.00	46.40	89.50
	others	0.00	0.00	0.00	0.00
	reversals	95.50	90.00	53.60	10.50
5000	x→y to y→x	4.80	3.20	0.10	0.00
	x→y to feedback	0.40	2.20	5.20	5.20
	x→y to non-caus	90.50	55.30	1.10	0.00
	x→y to x→y	4.30	39.30	93.60	94.80
	others	0.00	0.00	0.00	0.00
	reversals	95.70	60.70	6.40	5.20

Note: ρ gives the degree of cointegration (see the error correction term in the DGP); N is the sample size;→ : direction of causality, i.e., x→ y to y → x: x causes y in ECM to y causes x in VAR. "Others" means other cases, i.e., non-causality in ECM. "Reversals" is the frequency of change of x Granger-causing y in ECM into other relations in VAR. For each N and ρ, "reversals" is computed as the sum of rows (1), (2) and (3) divided by the sum of rows (1), (2), (3) and (4).

Table 2 contains the results of Granger causality reversals using the second set of parameters. As for the previous case, Granger causality distortion is a serious matter. For example, for a sample size of 100 with ρ=0.0, Granger causality reversal reaches 95 percent. The distortion depends on the degree of cointegration: the higher the degree of cointegration, the less likely is the distortion. Thus, for N=400, while Granger causality distortion reaches 92.8 percent for ρ=0.3, it is only 6.7 percent for ρ=0.9. Finally, the distortion tends to disappear as the sample size grows larger. Nevertheless, the speed of the disappearance of the distortion is very different between highly and weakly cointegrated variables. Thus, for example, for N=5000, while the distortion is only 5.4 percent for ρ=0.9, it is 96.3 percent for ρ=0.0. Overall, the two sets of parameters used in the DGP convey the same message about Granger causality distortion in a VAR in differences.

Table 2

Causality Pattern Reversals (in per cent) from ECM to VAR (see equations (9), (10) and (11) with 1000 replications using F statistics at the 5% level of significance:

DGP: $\alpha = -0.5, \beta_1 = 0.7, \beta_2 = 0.1, \rho = 0.0, 0.3, 0.6, 0.9$. $\rightarrow$: causality direction

N	Causality from ECM to VAR	$\rho = 0.0$	0.3	0.6	0.9
50	$x \rightarrow y$ to $y \rightarrow x$	2.80	2.90	2.70	3.00
	$x \rightarrow y$ to feedback	0.50	0.60	0.50	1.30
	$x \rightarrow y$ to non-caus	60.50	63.90	72.60	69.00
	$x \rightarrow y$ to $x \rightarrow y$	4.30	4.70	7.40	24.90
	others	31.90	27.90	16.80	1.80
	reversals	93.69	93.48	91.11	74.64
100	$x \rightarrow y$ to $y \rightarrow x$	5.20	5.90	6.20	2.50
	$x \rightarrow y$ to feedback	0.40	0.40	0.70	2.90
	$x \rightarrow y$ to non-caus	88.90	88.20	81.20	46.70
	$x \rightarrow y$ to $x \rightarrow y$	4.40	4.60	11.80	47.90
	others	1.10	0.90	0.10	0.00
	reversals	95.55	95.36	88.19	52.10

Contd.

Table : 2 (*Contd.*)

N	Causality from ECM to VAR	$\rho = 0.0$	0.3	0.6	0.9
200	x→y to y→x	4.40	4.00	4.70	0.80
	x→y to feedback	1.00	1.20	1.20	5.30
	x→y to non-caus.	89.80	88.50	75.50	15.20
	x→y to x→y	4.80	6.30	18.60	78.70
	others	0.00	0.00	0.00	0.00
	reversals	95.20	93.70	81.40	21.30
400	x→y to y→x	4.80	4.50	3.30	0.00
	x→y to feedback	0.80	0.70	2.10	5.50
	x→y to non-caus.	90.90	87.60	64.80	1.20
	x→y to x→y	3.50	7.20	29.80	93.30
	others	0.00	0.00	0.00	0.00
	reversals	96.50	92.80	70.20	6.70
1000	x→y to y→x	5.10	4.90	2.00	0.00
	x→y to feedback	0.20	1.20	4.20	6.50
	x→y to non-caus.	90.50	79.90	32.00	0.00
	x→y to x→y	4.20	14.00	61.80	93.50
	others	0.00	0.00	0.00	0.00
	reversals	95.80	86.00	38.20	6.50
5000	x→y to y→x	5.50	2.10	0.00	0.00
	x→y to feedback	0.50	3.30	5.10	5.40
	x→y to non-caus.	90.30	43.40	0.00	0.00
	x→y to x→y	3.70	51.20	94.90	94.60
	others	0.00	0.00	0.00	0.00
	reversals	96.30	48.80	5.10	5.40

Note: ρ gives the degree of cointegration (see the error correction term in the DGP); N is the sample size;→ : direction of causality, i.e., x→ y to y → x: x causes y in ECM to y causes x in VAR. "Others" means other cases, i.e., non-causality in ECM. "Reversals" is the frequency of change of x Granger-causing y in ECM into other relations in VAR. For each N and ρ, "reversals" is computed as the sum of rows (1), (2) and (3) divided by the sum of rows (1), (2), (3) and (4).

Concluding Remarks

This paper has attempted to answer three questions. First, is there any Granger causality distortion due to misspecifications (e.g., omission of an error correction term) in an unrestricted VAR in differences estimated with cointegrated variables? Second, is there any link between Granger causality distortion and the degree of "cointegratedness of variables"? Third, how does the sample size affect the answers to the two previous questions?

The paper reveals the following striking results. First, as expected an omission of an error correction term in unrestricted VAR in first differences translates into Granger causality distortions. Second, the seriousness of the latter distortions depends on the degree of cointegration of variables: the higher the degree of cointegration, the more likely is the distortion. Third, at each degree of cointegration, the larger the sample size, the less likely is the distortion. Fourth and more importantly, Granger causality distortion does decrease or eventually disappear in large samples. Nevertheless, the rate of disappearance of the distortion is substantially different between highly and weakly cointegrated variables.

Since the degree of cointegration of variables is not always checked in empirical studies and also since the distortion seems to persist in large samples with highly cointegrated variables, the paper strongly recommends the use of the ECM instead of a simple VAR even in large samples.

REFERENCES

Anderson, W., Wallace, M. and J.T. Warne (1986), "Government Spending and Taxation: What Causes What?," *Southern Economic Journal*, 52, 630–639.

Atkins, F.J. (1989), "Cointegration, Error Correction and the Fisher Effect," *Applied Economics*, 21, 1611–1620.

Banerjee, A. Dolado, J.J. Hendry, D.F., and G.W. Smith (1986), "Exploring Equilibrium Relationships in Econometrics Through Static Models: Some Monte Carlo Evidence," *Oxford Bulletin of Economics and Statistics*, 48, 253–277.

Campbell, J Y and R.J. Shiller (1988), "Interpreting Error Correction Mechanism," Journal of Economic Dynamics and Control, 12, 505–522.

Christiano, L.J and L Ljungqvist (1987), "Money does Granger-Cause Output in the Bivariate Output-Money Relation," Federal Reserve of Minneapolis, Staff Report 108

Darrat, A F. (1987), "Structural Deficits and Interest Rates: Some Causality and Cointegration Tests," *Southern Economic Journal*, 56, 752–758.

Davidson, J E H., Hendry, D.F, Srba, F and J S. Yeo (1978), "Econometric Modelling of the Aggregate Time SEries Relationship between Consumer's Expenditure and Income in the United Kingdom," *Economic Journal*, 88, 661–692

Engle, R.F and C.W J Granger (1987), "Cointegration and Error Correction. Representation, Estimation and Testing," Econometrica, 55, 251–276.

Engle, R.F. and B S. Yoo (1987), "Forecasting and Testing in Cointegrated Systems," *Journal of Econometrics*, 35, 143–159.

Garcia, P. and H O Zapata (1991), "Cointegration, Error Correction and The Fisher Effect: A Clarification," *Applied Economics*, 23, 1367–1368.

Geweke, J., Meese, R. and W. Dent (1983), "Comparing Alternative Tests of Causality in Temporal Systems· Analytic Results and Experimental Evidence," *Journal of Econometrics*, 21, 161–194

Granger, C.W J. (1969), "Investigating Relations by Econometric Models and Cross Spectral Methods,"Econometrica, 37, 428–438.

Granger, C.W.J. (1981), "Some Properties of Time Series Data and Their Use in Econometric Models Specification," *Journal of Econometrics*, 16, 121–130

Granger, C W J and A.A. Weiss (1983), "Time series Analysis of Error Correction Models," in Karlin, S et al.(eds), Studies in Econometrics, Time Series, and Multivariate Analysis, Academic Press, New York, 255–278.

Granger, C.W.J. (1983), "Co-Integrated Variables and Error Correcting Models," Discussion Paper, 83–13a, University of California, San Diego.

Hendry, D.F (1976), "The Structure of Simultaneous Equations Estimators," *Journal of Econometrics*, 4, 51–88

Hendry, D F. and J.F. Richard (1982), "On the Formulation of Empirical Models in Dynamic Econometrics," *Journal of Econometrics*, 20, 3–33

Hylleberg, S and G E. Mizon (1989), "Cointegration and Error Correction Mechanisms," *Economic Journal*, 99, 113–125

Hylleberg, S., Engle, R.F., Granger, C W.J. and B.S. Yoo (1990), "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215–238

MacDonald, R and C Kearny (1987), "On the Specification of Granger Causality Tests using Cointegration Methodology," *Economic Letters*, 25, 149–153.

Miller, S.M and Russek, F S. (1990), "Cointegration and Error Correction Models. The Temporal Causality between Government Taxes and Spending," *Southern Economic Journal*, 57, 221–229.

Mosconi, R and C Gianini (1992), "Non-Causality in Cointegrated Systems: Representation, Estimation and Testing," *Oxford Bulletin of Economics and Statistics*, 54, 399–417.

Nelson, C R. and C.I. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.

Ohanian, L.E , (1988), "The Spurious Effects of Unit Roots on Vector Autoregressions: A Monte Carlo Study," *Journal of Econometrics*, 31, 251–266.

Philips, A.W. (1954), "Stabilization Policy in a Closed Economy", Economic Journal, 64, 290–323.

Philips, P C.B. and S N Durlauf (1986), "Multiple Time Series with Integrated Variables," *Review of Economic Studies*, 53, 473–496.

Philips, P.C.B. and M. Loretan (1991), "Estimating Long Run Equilibrium," *Review of Economic Studies*, 58, 407–436

Sargan, J D. (1964), "Wages and Prices in the United Kingdom A Study in Econometric Methodology," in Hart, P E., Mills, G. and J.K Whitaker (eds.), Econometric Analysis for National Planning, Butterworths, London.

Sargent, T.J. (1976), "A Classical Econometric Model of The United States," *Journal of Political Economy*, 84, 207–237.

Sims, C A. (1972), "Money, Income and Causality," *American Economic Review*, 2, 540–552.

Sims, C.A. Stock, J.H and M.W.. Watson (1990), "Inference in Linear Time Series Models with Some Unit Roots,' *Econometrica*, 58, 113–144.

Spanos, A. (1990), "Unit Roots and Their Dependence on the Conditioning Information Set," in Fomby, T.B. and G F. Rhodes (eds.), Advances in Econometrics: Cointegration, Spurious Regression and Unit Roots, JAE Press, New York, 271–292.

Toda, Y. H and P.C.B. Phillips (1993), "The Spurious Effect of Unit Roots on Vector Autoregressions. An Analytical Study," *Journal of Econometrics*, 59, 229–255.

Wickens, M R. and T S. Breusch (1988), "Dynamic Specification, The Long Run and The Estimation of Transformed Regression Models," *Economic Journal* (conference), 98, 189–205.

Modelling Indirect Production with Expenditure Constraints: A Farm-Level Study of Rice Production in West Bengal

Anjana Samanta (Bhattacharyya)
Subal C. Kumbhakar

I. Introduction

The adequacy of productive resources is often an important question in the context of economic development, particularly for the developing countries facing high population growth. The growth of output in these countries, specially in the subsistence agricultural sector, is often limited by an acute shortage of productive inputs, viz. land, fixed capital equipment, and working capital required for the purchase of fertiliser and other inputs. Given this situation, the measurement of productivity of inputs and their substitutability seems to be a matter of great importance for these low-income countries. An enquiry into the nature of the production function might provide some guidelines for raising output. With this broad overall objective this paper seeks to measure the relationship among inputs and output in the context of rice cultivation in West Bengal.

Since the early sixties there has been widespread use of the production function in studies relating to Indian agriculture, primarily for measuring economic efficiency in the use of resources. These studies can broadly be classified into two groups depending on the type of approach used for analysing production technology. In the primal (direct) approach the production function was directly estimated, using the ordinary least squares technique. The production function was assumed to be Cobb-Douglas in most cases (Krishna 1963; Hopper 1965; Saini 1979; Sahota 1968; Chennareddy 1967; Sampath 1978; Ram 1980; among others). In the dual (indirect) approach, on the other hand, the cost or the profit function is studied to analyse the characteristics of the production function. The use of the dual approach gained immense popularity because of its several advantages. First, the dual functions are defined in terms of input and output prices, and the exogeneity of these variables allows for a consistent measurement of the dual function, even in terms of a single equation. Second, the use of the envelope theorem permits the system of input demand and output supply functions to be derived from the dual function by simple differentiation, facilitating the use of flexible functional forms. The estimation of the dual function along with these demand and supply equations increases the efficiency of parameter estimates, by taking into account the cross-equation restrictions on parameters as well as the covariance of the error terms. For all these reasons the later studies are based on the use of the dual approach (Lau and Yotopoulos 1971, 1972; Junankar 1980, 1982; Kalirajan 1981; Kumbhakar and Bhattacharyya 1992; Bhattacharyya and Glover 1993).

Despite these advantages of the dual approach the use of a particular dual formulation is justified only if the underlying behavioral hypothesis is appropriate. In the context of Indian agriculture, both profit and cost functions have some limitations. The use of the profit function, derived from profit-maximising behaviour, is implicitly based on the assumption that farms face no constraints in their input-allocation decisions. The validity of this assumption is highly questionable in the context of Indian agriculture. The use of the cost function is relatively more justified. In this formulation farmers are assumed to minimise the cost of a predetermined level of output. Although this approach incorporates some constraints on producers' behaviour, the exact nature of the constraint seems to be misspecified. In reality the large majority of producers of an underdeveloped economy are constrained by the availability of funds required for the purchase of variable inputs for a given agricultural season. This constraint is particularly binding for the small and marginal farmers who have very limited access to formal sector loans. In this situation rational behaviour on the part of the producers entails maximisation of output subject to a specified level of expenditure. This calls for the use of an alternative dual approach, the indirect production function approach, for the study of agricultural technology in the developing countries. In this paper we use a flexible indirect production function to analyse these properites. This methodology has not yet been directly used in applied production studies.

Apart from being a more appropriate characterization of producers' behaviour in underdeveloped countries the indirect production function approach has the advantage that it allows for the direct computation of the output or the budget effect (the production analogue of the income effect) resulting from a change in input price. The commonly used input demand elasticities, the partial elasticity of derived demand or the Allen-Uzawa elasticity of substitution, are designed to measure the response of input demand to factor price change for a constant level of output. An elasticity measure which takes into account the impact on output would often be of greater interest for public policy. In-fact the output effect of production studies is likely to be of considerable importance because of the limited number of inputs used in production process, particularly in agriculture. The importance of the output effect along these lines has been recognised by Chambers (1982). In his study the effect has been indirectly estimated from the cost function, using the inverse relationship between the cost function and the indirect production function. Field and Allen (1981) have considered a broader concept of output effect resulting from nonconstant output price in the presence of a downward sloping product demand curve, and also from nonconstant returns to scale. Although their variable-output input demand elasticity is a comprehensive measure of demand response to an input price change, it is of limited use in applied studies because of the inclusion of output demand elasticity in the elasticity expression for input demand. In this study we use the former concept of the output effect; the output change resulting from a change in input price in a constrained-budget situation.

The outline of the paper is as follows. The model presented in Section II, includes a brief discussion of the indirect production function in general terms. This is followed by a translog specification of the indirect production function and the derivation of the input cost share equations. Several different concepts of elasticity along with their properties, relative advantages and disadvantages are also discussed in this section. A brief description of the data set is given in Section III. The estimation procedure is discussed in Section IV. In Section V we analyse the empirical results, and Section VI presents the conclusions of this study.

II. The Model

a) The Preliminaries

The concept of the indirect production is relevant in the context of maximisation of output subject to a given technology, *m* quasi-fixed inputs and a given budget for the purchase of *n* variable inputs. Underlying this function, there is the familiar production function relating inputs and output

$$Y = f(X, Z) \tag{1}$$

where Y represents output, X denotes an $n \times 1$ vector of variable inputs, Z denotes the quasi-fixed input vector of order m. The resource (or budget) constraint faced by the farm can be written as

$$C = W'X \tag{2}$$

where W denotes the vector of hiring prices of variable inputs and C represents the fund available for the purchase of variable inputs. The producer wants to maximise output, given by (1), subject to the constraint in (2). The Lagrangean for the problem can be written as

$$\mathcal{L} = f(.) + \Lambda (C - W'X) \tag{3}$$

where Λ denotes the Lagrange multiplier associated with the constraint in (2). The exogenous variables are the elements in vectors Z, W and the total budget of the producer, C, while the input vector X and Λ are determined endogenously. Solving the first-order conditions for the problem, we get the solution of the endogenous variables in terms of the exogenous variables, viz.,

$$x_i = g_i(W,C,Z) \qquad \forall_i = 1,\dots,n \tag{4a}$$

$$\lambda = h(W,C,Z) \tag{4b}$$

Substituting the optimal values of x^1 from (4a) in (1) we get the optimal value of the objective function

$$y = \psi(W,C,Z) \tag{5}$$

---

[1] Here we have used upper case letters to denote the exogenous and endogenous variables, including vectors, while lower case letters denote the optimal values of the endogenous variables. All subscripts are in lower case letters.

Equation (5) represents the indirect production function (henceforth IPF). Given the technology, it expresses y, the maximum attainable output for the farmer in a specified period as a function of the availability of funds, the prices of variable inputs, and the amounts of fixed inputs.

Next we specify the functional form of the IPF. To impose minimum a priori restrictions on the underlying production technology we need a flexible functional form. The translog functional form is chosen because it imposes no a priori restrictions on any of the elasticities. With this specification the IPF can be written as

$$\ln y = \alpha_0 + \sum_{i=1}^{n} \alpha_1 \ln W_i + \alpha_c \ln C + \sum_{j=1}^{m} \phi_j \ln Z_j$$

$$+ \frac{1}{2}\left\{ \sum_{k=1}^{n} \sum_{i=1}^{n} \beta_{ik} \ln W_i \ln W_k + \beta_{cc} (\ln C)^2 + \sum_{l=1}^{m} \sum_{j=1}^{m} \mu_{jl} \ln Z_j \ln Z_l \right\}$$

$$+ \sum_{j=1}^{m} \sum_{i=1}^{m} \gamma_{ij} \ln W_i \ln Z_j + \sum_{i=1}^{n} \gamma_{ic} \ln W_i \ln C + \sum_{j=1}^{m} \theta_{jc} \ln Z_j \ln C. \quad (6)$$

In addition to the usual symmetry restrictions on the coefficients, β_{ik}, μ_{jl}, γ_{ij}, economic theory tells us that the IPF is homogeneous of degree zero in input prices and C. This gives rise to the following set of restrictions on the parameters of the model

$$\sum_{i=1}^{n} \alpha_i + \alpha_c = 0 \qquad (7)$$

$$\sum_{k=1}^{n} \beta_{ik} + \gamma_{ic} = 0 \qquad \forall_i = 1,\ldots,n, \qquad (8)$$

$$\sum_{i=1}^{n} \gamma_{ij} + \theta_{jc} = 0 \qquad \forall j = 1,\ldots,m, \qquad (9)$$

$$\sum_{i=1}^{n}\gamma_{ic}+\beta_{cc}=0 \qquad (10)$$

From the IPF we can derive the constant-cost demand function[2] for the ith input using Roy's identity[3]. This gives us

$$x_i = -\frac{\dfrac{\partial y}{\partial W_i}}{\dfrac{\partial y}{\partial C}} \qquad (11)$$

This enable us to determine the share of the ith input in total variable cost. It is given by

$$S_i = -\frac{\dfrac{\partial \ln y}{\partial \ln W_i}}{\dfrac{\partial \ln y}{\partial \ln C}} = -\frac{e_{yi}}{e_{yc}} \qquad (12)$$

where e_{yc} denotes the elasticity of output with respect to a change in the producer's budget and e_{yi} denotes the elasticity of output with respect to a change in the price of the *i*th variable factor. This brings us to the discussion of elasticities.

b) Elasticities

In this study we will consider several different concepts of elasticity. The ordinary elasticities of output supply and input demand for our constrained–budget situation, known as the constant–cost elasticities, will be considered first. The output supply elasticities for changes in the farmer's budget, and also for input price changes, have already been introduced. For our model the expressions for these elasticities are

---

[2] We refer to this input demand as constant-cost demand to distinguish it from the constant-output demands attainable from cost-minimisation problems.

[3] The Roy's Identity is an application of the envelope theorem which enables us to determine the change in the optimal value of the objective function for a marginal change in the exogenous variables. For the present maximisation problem the use of this theorem and the substitution of the value of $\partial y/\partial C$ in the expression for $\partial y/\partial W_i$ gives rise to the Roy's identity.

$$e_{yc} = \frac{\partial \ln y}{\partial \ln C} = \alpha_c + \beta_{cc} \ln C + \sum_{i=1}^{n} \gamma_{ic} \ln W_i + \sum_{j=1}^{m} \theta_{jc} \ln Z_j \qquad (13)$$

$$e_{yi} = \frac{\partial \ln y}{\partial \ln W_i} = \alpha_i + \gamma_{cc} \ln C + \sum_{k=1}^{n} \beta_{ik} \ln W_k + \sum_{j=1}^{m} \gamma_{ij} \ln Z_j \qquad (14)$$

In general, the sign of e_{yc} is positive indicating an increase in optimal output resulting from a relaxation of the farmer's budget constraint. Similarly, e_{yi} is in general negative indicating a fall in output caused by a reduction of the budget in real terms, following a rise in input price.

Next we consider the effect of a change in the *j*th quasi-fixed input on optimal output in terms of the following elasticities, viz.,

$$\eta_{yj} = \frac{\partial \ln y}{\partial \ln Z_j} = \phi_j + \sum_{l=1}^{m} \mu_{lj} \ln Z_l + \sum_{i=1}^{n} \gamma_{ij} \ln W_i + \theta_{jc} \ln C, \; j=1,\ldots,m \qquad (15)$$

A positive value of this elasticity indicates a rise in output with an increase in the amount of the *j*th quasi-fixed input. This is because of the positive marginal product of the relevant input.

To determine the elasticities of input demand let us rewrite equation (12) in the form

$$\ln X_i = \ln C - \ln W_i + \ln\left\{-\frac{e_{yi}}{e_{yc}}\right\} \qquad (16)$$

The elasticities of constant-cost input demand can be derived from this equation. The own-price elasticity for the *i*th input is

$$e_{ii} = \frac{\partial \ln X_i}{\partial \ln W_i} = \frac{\beta_{ii}}{e_{yi}} - \frac{\gamma_{ic}}{e_{yc}} - 1 \qquad (17)$$

This is expected to have a negative sign, unless there is a strong offsetting budget effect. The cross-price elasticity, the sign of which determines the presence or absence of substitutability, is given by

$$e_{ik} = \frac{\partial \ln X_i}{\partial \ln W_k} = \frac{\beta_{ik}}{e_{yi}} - \frac{\theta_{kc}}{e_{yc}}, \quad i \neq k \tag{18}$$

For changes in the producer's budget the impact on the demand for the *i*th variable input is measured by the following elasticities

$$e_{ic} = \frac{\partial \ln X_i}{\partial \ln C} = 1 + \frac{\gamma_{ic}}{e_{yi}} - \frac{\beta_{cc}}{e_{yc}} \tag{19}$$

A positive value of this elasticity indicates that the *i*th input is a normal one, while a negative value would classify the input as inferior.

The elasticity of demand for the *i*th variable input for a change in the amount of the *j*th quasi-fixed input is given by

$$\eta_{ij} = \frac{\partial \ln X_i}{\partial \ln Z_j} = \frac{\gamma_{ij}}{e_{yi}} - \frac{\theta_{jc}}{e_{yc}} \tag{20}$$

This elasticity, known as the elasticity of intensity (Diewert 1974; Dupont 1991), can have a positive or negative value, depending on the nature of substitutability between the pairs of variable and quasi-fixed inputs.

Next we consider some other concepts of elasticity. It is well-known that the twin problems of output-maximisation in a situation of budget constraint and of cost-minimisation for a given level of output are mirror images of one another. They generate the same input demand vector if the value of the output parameter for the second problem is set at the solution value of output for the former. From this equality it is possible to derive the relationship between the constant-cost elasticity of demand for the *i*th input for a change in the *k*th factor price and the corresponding constant-output elasticity. This is given by the production counterpart of the Slutsky equation, written in elasticity form (Chambers 1982).

$$e_{ik} = \epsilon_{ik} - e_{ic} S_k \tag{21}$$

where ϵ_{ik} denotes the elasticity of constant-output input demand. The second term on the RHS shows the budget effect (the production analogue of income effect). It captures the effect of a change in the *k*th factor price on the demand for the *i* th factor through a change in the producer's purchasing power over inputs.

In standard production analyses elasticities, ϵ_{ik} are obtained by applying Shephard's Lemma[4] to the cost function. Hence the properties of the constant-output demand functions reflect those of the cost function. The concavity of the cost function in W implies that the own-price elasticities, for the constant-output case, must be negative. This is not true for constant-cost own price elasticities. It is possible for such an elasticity to be positive only if e_k is negative, i.e., the input is an inferior input, and the budget effect is stronger than the pure substitution effect.

The classification of inputs as substitutes or complements depends, obviously, on the sign of the cross-price elasticities. If e_{ik} is positive then input *i* is a substitute for input *k*, otherwise it is a complement. Similar remarks hold for ϵ_{ik}. Now the question is one of symmetry. If input *i* is found to be a substitute for input *k*, does it necessarily follow that the *k*th input will also be a substitute for the *i* th one? The answer depends on the type of elasticity being considered. Symmetry holds for constant-output elasticity, ϵ_{ik}, but not for the constant-cost elasticity, e_{ik}. Although the constant-output elasticity ϵ_{ik} is not equal in value to ϵ_{ki}, they have the same sign. This follows from the properties of the cost function. On the other hand e_{ik} is symmetric neither in sign nor in value. This is because of the presence of the budget effect in the expression for this elasticity. There is no reason for the budget effects in the two expressions to have the same sign, let alone the same value. Hence for cross-elasticities measured along an isoquant in *n*-space the classification of inputs as substitutes or complements remains invariant to the direction in which we are moving in response to a price change, but it is not so for input demand change when adjustments in output are allowed for.

---

[4] This is another application of the envelope theorem in the context of the cost minimisation problem to derive the input demand function.

For a meaningful discussion of input substitutability it is necessary to go beyond the simple own-and cross-price elasticity measures. The concept of elasticity of substitution has been devised to provide more insight. The most widely used measure, the Hicks-Allen-Uzawa elasticity of substitution, is designed to measure the percentage change in input ratio following a one percent change in the marginal rate of technical substitution (MRTS) along a given isoquant. This elasticity is claimed to reflect the curvature properties of the isoquant. Under perfect competition the MRTS is equated to the input price ratio, hence the elasticity is supposed to represent the ratio of the percentage change in input ratio to the percentage change in their relative price. In the two-inputs case the elasticity of substitution formula, as derived by Hicks, truly satisfies this property. But in the presence of many inputs, the Allen-Uzawa generalisation of the elasticity expression does not adequately reflect the response. As mentioned by Chambers (1988) a change in the ratio of constant-output demands for the ith and kth inputs, denoted by x_i^* (W, Y) and x_k^* (W,Y) respectively, may be brought about in several ways. It may be caused by a change in x_i^* alone, or by a change in x_k^*, or because of simultaneous changes in both variables. A measure of a change in the input ratio should be capable of capturing all these changes. This is not satisfied by the Allen-Uzawa measure.

As an indicator of factor substitutability, in relative sense, the expression for the elasticity of substitution should be of the form $\left(\hat{x}_i^* - \hat{x}_k^*\right) \big/ \left(\hat{W}_k - \hat{W}_i\right)$, where $\hat{x}_i^*$ represents the proportionate change in the variable x_i^* and so on. Only in such cases can it be termed a real two-factor-two-price elasticity of substitution (TTES). However, the expression for the Allen-Uzawa elasticity in the multiple-inputs case can be reduced to the form

$$\sigma_{ik} = \frac{\epsilon_{ik}}{S_k} \qquad (22)$$

This indicates that the Allen-Uzawa elasticity of substitution provides no more information than ϵ_{ik}, it is simply the constant-output cross elasticity, divided by the cost-share of the input whose price has changed. As W_k changes the Allen-Uzawa elasticity simply measures the change in the input ratio x_i^*/x_k^* brought about by a change in the numerator, but fails to capture the change in the ratio because of an adjustment in the use of the *k*th input following a change in its own price. Hence there seems to be little justification for preferring the Allen-Uzawa elasticity of substitution to the simple constant-output elasticity. Like ϵ_{ik} it is a simple one-factor-one-price elasticity of substitution (OOES). The ony convenience of σ_{ik} is that it is symmetric in value, which is not true for ϵ_{ik}.

In this context the study of the Morishima elasticity of substitution seems to be more justified (Blackorby and Russel 1989). The Morishima elasticity of substitution can be expressed as

$$\rho_{ik} = \epsilon_{ik} - \epsilon_{kk} \tag{23}$$

From this equation it is obvious that Morishima's measure is a two-input-one-price elasticity of substitution (TOES) because it is of the form $\left(\hat{x}_i^* - \hat{x}_k^*\right) / \hat{W}_k$. It captures the proportionate change in x_i^* as well as x_k^* as a result of the change in W_k, and so it provides more information than the Allen-Uzawa elasticity of substitution.

Again the question of symmetry is relevant. From equation (23) it can be seen that $\rho_{ik} \neq \rho_{ki}$ because the various elasticities included in the two expressions may not be equal.As to the symmetry of sign an interesting observation can be made. For this we need to take a closer look at the elasticity expression (23) and the expression given below

$$\rho_{ki} = \epsilon_{ki} - \epsilon_{ii} \tag{24}$$

If the constant-output cross-elasticities for inputs *i* and *k* are positive, so that they are substitutes along an isoquant,then they are Morishima substitutes as well.This is because the second term on the RHS, including the minus sign, is positive in both cases. So the sum can only be positive for both expressions. This, however, is not true if the relevant constant-output cross elasticities are negative, i.e., the inputs are complements. In this case it is possible for one or both of the Morishima elasticities to be positive, as the second term may outweigh the first term. Hence if input *i* is a Morishima complement for input *k* it does not follow that k will be a complement for i. It all depends on the relative magnitudes of the two elasticities involved in the relevant equation.

Thus the classification of inputs as substitutes or complements is not invariant to the choice of the elasticity of substitution formula. If the constant-output elasticities for inputs i and k are negative, then the Allen-Uzawa measure would also classify them as complements, while by the Morishima measure input i can be a substitute for input k. The economic intuition is as follows. The Allen-Uzawa elasticity simply tells us that the constant-output demand for i th input falls as the price of the kth one rises. But it is also true that the demand for the kth input falls with a rise in its own price. If the proportionate fall in x_k^* is greater than the proportionate diminution in x_i^* then the ratio of x_i^* will increase, indicating input i to be a substitute for input k, in a relative sense. However, if two inputs are substitutes by the Allen-Uzawa measure then they are Morishima substitutes as well. In this case both the rise in demand for input i and the fall in demand for input k work in the same direction, raising the ratio of i to input k. Hence we get similar classification by both measures, although the magnitudes of the effects are different.

Now one may ask the question: which one is more suitable?. There is no simple answer to this question. The two measures relate to two different concepts, and so the choice depends on the purpose of investigation. If we seek to estimate a simple change in the absolute demand for the *i*th input for a change in the *k*th input price then the knowledge of the elasticity ε_{ik} will suffice. In fact no additional information is provided by the Allen-Uzawa measure. On the other hand, if the objective is to measure the change in the input ratio then the relevant concept is the Morishima elasticity. Thus for more than two inputs the Allen-Uzawa elastictity is not a valid concept of elasticity of substitution, as the concept was originally introduced. Also it is redundant when the constant-output elasticity is already known. If there is a change in the price of only one input, W_k then the Morishima elasticity, ρ_{ik}, provides all relevant information regarding the adjustment in the input use ratio in response to the change in the ratio of factor, W_i / W_k. However, if there is a simultaneous change in the *i*th price as well, the Morishima measure would be inadequate, and it would be necessary to compute a TTES, for analysing the net impact on the input ratio.

A familiar measure of TTES between two inputs, evaluated along the factor-price-frontier for a constant level of expenditure on variable inputs, is known as the shadow elasticity of substitution. The expression for the shadow elasticity between inputs i and k is given by

$$K_{ik} = \frac{S_i}{S_i+S_k}\rho_{ik} + \frac{S_k}{S_i+S_k}\rho_{ki} \qquad (25)$$

From the above expression it can be seen that the shadow elasticities are symmetric. They capture the total effect of a change in the relative price of two inputs on their relative use.

III. Data Description

The data set used in this study has been collected by the Agro-Economic Research Centre on behalf of the Ministry of Agriculture of India. It consists of a sample of 289 farm households, spread over three agro-climatic zones of West Bengal. Region I covers two districts, Nadia and Murshidabad. Region II covers the districts of Burdwan and Birbhum, and the third region includes four districts, viz. Howrah, Hooghly, 24 Parganas and Midnapore. Together these three regions cover almost the entire paddy-producing belt of West Bengal. The data from these households have been collected over the period 1980–1985. It is not a panel data, the regions were surveyed in different years. For our study we have used the data on aman paddy, the principal crop of the state.

The survey provides a wide range of information on the fixed and variable inputs used for paddy cultivation. For a given agricultural season the amounts of land and capital can be assumed to be given for a farm household, hence these two inputs are treated as quasi-fixed inputs. In general, the quantities of human labour, bullock labour and fertiliser can be treated as endogenous variables for a given agricultural season. In the context of Indian agriculture a large part of human labour is supplied by members of the farm household.

Here it is necessary to make a distinction between family and hired labour, because the amount of working capital available for the purchase of variable inputs from the market is the main constraining factor in our analysis. Keeping in mind the requirements of the model we treat only hired labour as variable input. The other component of human labour, family labour, is considered as a quasi-fixed input. In terms of our model the implication of this assumption is: the decision with respect to the use of family labour is independent of the market wage rate, and the amount of family labour does not make a drain on the limited financial resources available for the purchase of variable inputs. Hence in our study capital (K), land (L) and family labour (M) are the quasi-fixed factors, and the variable inputs are hired labour (N), bullock labour (B), and fertiliser (F).

The land variable represents the amount of land under cultivation, measured in acres. Capital stock, by definition, is the total value of all durable farm assets, viz. the farm house, irrigation and all sorts of farm equipment. However, in applied production analysis it is not necessary to take into account the entire stock of capital, but the flow of services from capital. This brings us to the concept of the user cost of capital, defined as the sum of the market interest rate and the yearly depreciation rate, minus the yearly inflation rate (Jorgenson 1967). In our analysis the user cost is computed, using the data on average market interest rate, inflation rate and assuming a straight line depreciation at the rate of 10%.

Human labour, hired as well as family labour, is measured in man-days (eight hours per day.) Family labour represents the quantity of household labour used for cultivation during a particular agricultural season, and hired labour indicates the quantity hired from the market. The ruling market wage rate is used as the price of hired labour.

The fertiliser input represents the amount of chemical fertiliser used for paddy cultivation, measured in kilograms. The fertiliser market was highly regulated during the survey period. As a consequence a parallel market exists and the farmers have free access to that market. In our study the actual price of fertiliser rather than the controlled price is used as the effective price of the input.

The other variable inputs is bullock labour. It is measured in bullock-labour days (eight hours per day). Its price is the rental price for a pair of bullocks (per day). Output is the gross production of paddy, measured in quintals.

Estimation of the dual functional form requires variations in input prices. For the data set used in this study input prices show considerable variations, not only across regions but also within a particular region. These variations can be attributed, at least in part, to the differences in transportation costs and quality. But a more significant determinant of price seems to be the nature of forward contracts that a farmer enters into with local input sellers and buyers of paddy. Often inputs are sold at a different price as a part of the contract for output sale, an example of interlinked contracts. Although the Minimum Wage Law was in effect during the survey period, wide variations in wage rates are observed.

IV. Method of Estimation

The econometric model in the present setup consists of the indirect production function and the (n–1) cost share equations. After appending a classical random disturbance term to each of the equations we get the following set of equations for estimation

$$\ln y_f = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln W_{if} + \alpha_c \ln C_f + \sum_{j=1}^{m} \phi_j \ln Z_{jf} + \frac{1}{2}$$

$$\left\{ \sum_{k=1}^{n} \sum_{i=1}^{n} \beta_{ik} \ln W_{if} \ln W_{kf} + \beta_{cc} (\ln C_f)^2 + \sum_{l=1}^{m} \sum_{j=1}^{m} \mu_{jl} \ln Z_{jf} \ln Z_{lf} \right\}$$

$$+ \sum_{j=1}^{m} \sum_{i=1}^{n} \gamma_{ij} \ln W_{if} \ln Z_{jf} + \sum_{i=1}^{n} \gamma_{ic} \ln W_{if} \ln C_f + \sum_{j=1}^{m} \theta_{jc} \ln Z_{jf} \ln C_f + u_f \tag{26}$$

$$S_{if} = -\frac{e_{yif}}{e_{ycf}} + \xi_{if} \tag{27}$$

where

$$e_{ycf} = \alpha_c + \beta_{cc} \ln C_f + \sum_{i=1}^{n} \gamma_{ic} \ln W_{if} + \sum_{j=1}^{m} \theta_{jc} \ln Z_{jf} \tag{28}$$

$$e_{yif} = \alpha_i + \gamma_{ic} \ln C_f + \sum_{k=1}^{n} \beta_{ik} \ln W_{kf} + \sum_{j=1}^{m} \gamma_{ij} \ln Z_{jf} \quad (29)$$

and f indexes farms (f=1, ..., 289). The error vector (u_f, x_{Bf}, x_{Nf}) is assumed to have zero mean and constant variance-covariance matrix, independent across farms. Since the three input share equations are not independent (their sum being unity), one of the share equations has to be dropped to avoid the problem of singularity of the disturbance covariance matrix for these equations. The estimates are, however, invariant to the choice of the share equation to be dropped. Here we dropped the cost share equation for fertiliser. Thus we have a system of nonlinear seemingly unrelated regression (NLSUR) equations where the error terms are correlated across equations. We used an iterative procedure for estimation. The estimators thus obtained are consistent and asymptotically normal (Amemiya 1986). It is to be noted that the multi-variate normality assumption for the error vector is not used.

Three separate models are estimated. First, we estimate the coefficients and the elasticities for the entire sample of 289 farm households (pooled model). Then the model is estimated separately for large and small farms[5]. The likelihood ratio test is used to check for a significant divergence between the parameter estimates from the two size-classes. The test result shows that the coefficients are different for the two groups, indicating the need for separate estimation[6].

V. Analysis of Results

This section summarizes the results obtained from the pooled model, as well as the ones by size-class. Various elasticities, as mentioned in Section II.b, are estimated. Table 1 reports the elasticities for all models, evaluated at the mean values of the data. Since the elasticities are mostly nonlinear functions of the parameters and data, their standard errors are obtained by taking a linear approximation of these functions. The parameter estimates, along with their asymptotic standard errors, from all the models are reported in the Appendix.

---

[5] In this study the farms holding more than five acres of land are considered to be large, while those with smaller land-size are classified as small. According to this definition 110 farms are large.

[6] The Likelihood Ratio test statistic, $[-2(L_R-L_U)]$, is found to be 246.47 which exceeds the critical value of χ^2 (with 26 degrees of freedom) at any reasonable level of significance thereby rejecting the restricted model. L_R and L_U are the values of the log likelihood functions corresponding to the restricted (pooled) and unrestricted (separate models for small and large farms) models, respectively

Of prime importance from the standpoint of our methodology are the elasticities of output supply and input demand, for a marginal change in the producers' budget, i.e., e_{yc} and e_{ic}. It is to be noted here that all elasticities are farm-specific, i.e. there are 289 separate values for each elasticity. Since reporting all the elasticities for each data point is not practical, we report their estimates at the mean values of the data. The results indicate that, for the combined sample, output is expected to rise by 0.34 percent for a one percent rise in budget. For large and small farms these values are 0.15 and 0.28, respectively. For hired labour the demand is likely to rise more than proportionately, for all size-classes, with a change in the producers' budget. For a one percent rise in budget increases in demand for hired labour are in the order of 1.38 percent for all farms combined, 1.58 percent for large farms, and still higher for small farms, approximately 1.65 percent. The corresponding figures for bullock labour are 0.57, 0.26 and 0.32 percent, respectively. For fertiliser the respective figures are 0.85, 0.7 and 0.83 percent. All these elasticities are found to be highly significant. Our results also highlight the importance of the cost constraint: the higher responsiveness of small farms to budget changes indicates their greater need for working capital, compared to their larger counterpart.

The amounts of quasi-fixed inputs, capital and land, also impose some restrictions on attainable output and thereby determine the levels of demand for the variable inputs. The output supply elasticity, e_{yj}, measures the expected change in output for a marginal increase in the amount of the jth quasi-fixed input. Our results indicate that a one percent rise in capital stock will raise output by 0.11 percent for the pooled sample. For the groups of large and small farms the respective figures are 0.118 percent and 0.124

percent. For a similar change in land size output is expected to grow by 0.48 percent in the pooled model. The output increases for the large and small farms are 0.64 percent and 0.84 percent, respectively. In comparison, an increase in the quantity of family labour will have a smaller impact on output, 0.08 percent for the small farms, and 0.03 percent in the remaining cases. This seems to indicate overutilisation of family labour in all farm households, irrespective of land size, resulting in low marginal product of this input. For each hired input the increase in price causes a further stringency in the budget situation and consequently lowers output. A doubling of the wage of hired labour is expected to bring about an output reduction of 17 percent for the pooled sample, 8 percent for large farms, and 12.9 percent for the small farms. Similar effects, of a somewhat smaller magnitude, are observed for the rise in the prices of bullock labour and fertiliser. Once again the output effects are more pronounced for the small farms. This is in line with their more restrictive budget situation, making them more vulnerable to market price changes. All these effects are, once again, statistically significant.

Demand for the varible inputs can, in general, change in either direction following an increase in the amount of a particular quasi-fixed input used. In our analysis broadly same pattern of relationship is observed, with some differences in the elasticity values, for all three models. Hired labour and capital are found to be substitutes, as evidenced by a fall in labour demand with a rise in the stock of capital. In contrast, the demands for bullock labour and fertiliser are found to increase with rise in capital stock, hence these two inputs can be clasified as complements of capital. A rise in land size is accompained by an increase in the demand for bullock labour, and a decline in demand for hired labour as well as fertiliser. An increase in the amount of family labour brings about a substitution of hired labour, as expected. The resulting cost-saving allows them to purchase more of the other variable inputs, fertiliser and bullock labour.

The own price elasticities of demand (e_{ii}) for all variable inputs are negative, for the constant-output situation. This negativity follows from the properties of the cost function. The magnitudes of these elasticities, except for bullock labour, are higher for small farms. As to the constant-cost case, all such elasticities (e_{ii}) are negative for the pooled sample as well as the sub-samples of small and large farms.

Finally, we come to the question of substitutability among the variable inputs. The simplest measure is provided by the cross price elasticities. The constant-output elasticities (e_{ik}) indicate the same broad pattern for the combined sample and the sample of small farms. In both groups all three variable inputs are observed to be substitutes for one another. For large farms, while there is substitutability between hired labour and fertiliser, bullock labour and fertiliser are found to be complements, so are human and bullock labour. The elasticities showing complementarity are, however, statistically insignificant. The constant-cost cross-price elasticities (e_{ik}) indicate the same type of relationship for all three samples. Human and bullock labour are complements of one another, so are bullocks and fertiliser. With respect to human labour and fertiliser, however, the relationship is not reciprocal. As the price of fertiliser rises the demand for labour decreases, indicating complementarity, but for a rise in the price of labour the relationship is one of substitutability. While all variable inputs appear as substitutes in the production of a given level of output, at least for the small farms and the pooled sample, the dominant relationship is one of complementarity for a constant level of expenditure on variable inputs. Once again this indicates the importance of the output effect. With an increase in the price of an input, say hired labour, the demand for bullock labour is expected to rise, if output were to remain constant; but actually the price increase lowers output and the demand for bullock labour by causing a drain on the farmers' limited resources, due to the high budget share of human labour input. Hence the constant-cost elasticity for these two inputs turns out to be negative. In general, for analysing the effect of an input price change on the quantities of other inputs, what matters is the total effect, rather than the pure substitution effect.

The Allen-Uzawa elasticities of substitution (σ_{ik}) present the same picture as the cross-elasticities, for reasons discussed in

Section IIb. The Morishima elasticities (ρ_{ik}) provide more insight by measuring the response of the input use ratio to an input price change. Here, once again, the same pattern of relationship is observed for the small farms and the pooled sample: all variable inputs are found to be substitutes for one another. For large farms there is reciprocal complementarity between human and bullock labour, fertiliser and human labour are substitutes for one another. The reciprocity is not observed for bullock labour and fertiliser. A doubling of fertiliser price induces a 29 percent rise in the ratio of bullock labour to fertiliser, indicating substitutability between these two inputs, while a similar change in the price of bullock labour brings about a 39 percent reduction in the relative use of fertiliser, giving rise to the notion of complementarity. The lack of symmetry can be attributed to the difference in the relative magnitudes of the effects on the two inputs. Although the absolute use of bullock labour decreases with a rise in fertiliser price, the demand for the input increases in relative terms, because of a strong negative own price effect on the demand for fertiliser. On the other hand, for a rise in the price of bullock labour, the own price effect is not strong enough to outweight the negative impact on fertiliser use, causing a relative as well as absolute decline in the consumption of fertiliser.

Finally, we come to the show elasticities (κik) which present a comprehensive picture of relative substitution among inputs. For the pooled sample and small farms, once again, there is substitutability among all variable inputs, with statistically significant values of all elasticities. For the pooled model a one percent rise in the price of hired labour relative to the price of fertiliser brings about a 0.77 percent increase in the use of fertiliser relative to hired labour. For large and small farms the input ratio changes are in the order of 0.45 and 0.63 respectively. Between bullock labour and fertiliser the relationship is one of substitutability for all three cases. A doubling of the price ratio brings about an increase in the relative use of the relatively cheaper input, by 62 percent for the pooled data, 17 percent for large farms, and by 46 percent for the small ones. For a similar change in the price ratio between human and bullock labour the relative use of the cheaper input (for which the relative price has declined) increases by 0.47 percent for the pooled model,and by 0.31 percent for small farms. Substitutability between these two inputs is not observed in the case of large farms: for a doubling of the relative price of human labour its relative use is found to increase by 22 percent. This effect, however, is not significant. On the whole, the predominant form of relationship among the variable inputs is one of substitutability, particularly for the small farms and the aggregate sample. Once again the small farms exhibit greater response to relative price changes.

VI. Conclusions

This paper focuses on some problems of the subsistence agricultural sectors of the developing countries which have not received proper attention. In addition to the well-known constraints of small holding size and scarcity of capital equipment, the limited amount of working capital available for the purchase of variable inputs is found to be an important factor limiting the growth of output. This is revealed by the high elasticities of output supply and input demand for a marginal relaxation of the budget constraints of the farmers, for all size classes of farms. The elasticities are relatively larger for the small farmers, indicating their greater need for production loans. In the Indian situation this seems to call for a rethinking on the lending policies of the nationalised banking sector. A policy of making formal sector loans more accessible to the small and marginal farmers seems to be desirable in the interest of raising agricultural output in a resource-constrained economy with high population growth.

Another interesting finding of our study relates to the relative magnitudes of the output-supply elasticities for increases in the quantities of the quasi-fixed inputs. For increases in the two main fixed inputs, land and capital, output growth is relatively higher for the small farms. This seems to indicate the potential benefits from a reallocation of these fixed inputs among different classes of farm households. On the whole, this study highlights the importance of the resource constraints of the small and marginal farmers which deserve special attention for plicy formulation in developing countries.

REFERENCES

Amemiya, T., *Advanced Econometrics*, Harvard University Press, Cambridge, 1986

Bhattacharyya, A., and T F. Glover., "Profit Inefficiency of Indian Rice Farmers A System Approach." *Journal of Productivity Analysis* 4(1993): forthcoming.

Blackorby, C and R R. Russell, "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)" *American Economic Review* 79(1989).882–88

Chambers, R G., "Duality, the Output Effect, and Applied Comparative Statics" *American Journal of Agricultural Economics* 64(1982): 152–156

——, *Applied Production Analysis. A Dual Approach.* Cambridge. Cambridge University Press, 1988.

Chennareddy, V., "Production Efficiency in South Indian Agriculture." *Journal of Farm Economics* 49(1967).

Diewert, W.E., "Applications of Duality Theory" *Frontiers in Quantitative Analysis*, ed. M.D Intriligator and D.A. Kendrick. Amsterdam: North-Holland Publishing Co., 1974.

Dupont, D P., "Testing for Input Substitution in a Regulated Fishery." *American Journal of Agricultural Economics* (1991): 155–164.

Field, B C , and P.G. Allen., "A General Measure for Output-Variable Input Demand Elasticities" *American Journal of Agricultural Economics* 63 (August 1981).

Hopper, D W , "Allocative Efficiency in Traditional Indian Agriculture" *Journal of Farm Economics* 47(1965): 611–624.

Jorgenson, D.W., "The Theory of Investment Behavior" in *Determinants of Investment Behavior*, National Bureau of Economic Research, New York, 1967.

Junankar, P.N., "Do Indian Farmers Maximise Profits?" *Journal of Development studies* 16(1980): 187–203.

——, "Neo-Classical Economics and Indian Agriculture: An Econometric Analysis of Producer Behaviour" *Indian Economic Review* 17(1982): 139–164.

Kalirajan, K., "The Economic Efficiency of Farmers Growing High Yielding Irrigated Rice in India" *American Journal of Agricultural Economics* 63(1981): 566–570.

Krishna, R , "Some Production Functions for Punjab." *Indian Journal of Agricultural Economics* 19(July-Dec 1964)

Kumbhakar, S.C., and A Bhattacharyya, "Price Distortions and Resource-use Efficiency in Indian Agriculture. A Restricted Profit Function Approach." *The Review of Economics and Statistics* LXXIV (1992) 231–239.

Lau, L.J , and P A Yotopoulos., "Profit, Supply and Factor Demand Functions." *American Journal of Agricultural Economics* 54(1972) 11–18.

——, "A Test of Relative Efficiency and Application to Indian Agriculture " *American Economic Review* 61(1971)·94–109.

Ram, R , "Role of Education in Production: A Slightly New Approach."*Quarterly Journal of Economics* 95(1980): 365–373.

Sahota, G.S , "Efficiency in Resource Allocation in Indian Agriculture." *American Journal of Agricultural Economics* 50(1968). 584–605.

Saini, G R , *Farm Size, Resource-Use Efficiency and Income Distribution*, Allied Publishers Private Limited, New Delhi, 1979.

Sampath, R K , *Economic Efficiency in Indian Agriculture-Theory and Measurement* India. Macmillan, 1978.

Table 1

Estimates of Elasticities

	Pooled Model		Large Farms		Small Farms	
Elasticities	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Output supply elasticities with respect to cost and variable inputs						
e_{yc}	.3426	.0238	1512	.0333	2771	.0333
e_{yN}	-.1701	.0123	- 0799	.0177	-.1293	.0162
e_{yB}	- 1359	.0106	- 0573	.0126	-.1164	.0145
e_{yF}	- 0367	0031	-.0140	0032	-.0315	.0042
Input demand elasticities with respect to cost						
e_{Nc}	1.3765	.0309	1.5850	.1306	1.6533	.0834
e_{Bc}	.5679	0314	.2585	.1621	.3200	.0833
e_{Fc}	.8545	0152	.6967	.0670	.8312	0201
Output supply elasticities with respect to quasi-fixed inputs						
η_{yK}	.1083	.0169	.1181	.0231	1245	0234
η_{yL}	.4804	.0646	6421	.1191	8422	1857
η_{yM}	.0332	.0163	0325	0158	.0770	.0283

Contd.

Table . 1 (*Contd.*)

	Pooled Model		Large Farms		Small Farms	
Elasticities	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Input demand elasticities with respect to quasi-fixed inputs						
η_{NK}	- 0337	0262	- 1067	0492	-.0342	.0429
η_{NL}	-.0011	.0894	- 2794	2071	- 8642	.3303
η_{NM}	-.0174	.0223	- 0336	.0367	- 1119	0537
η_{BK}	.0273	.0387	.0774	0496	0336	.0510
η_{BL}	.0553	.1093	4241	.2852	1 1211	3120
η_{BM}	.0090	.0243	0273	.0339	0819	0451
η_{FK}	.0551	.0411	2564	.0794	0160	.0629
η_{FL}	- 1999	.1353	-.1396	.3155	- 5958	4527
η_{FM}	0476	0289	0800	0482	.1569	0642
Constant-output input demand elasticities (own and cross)						
ε_{NN}	- 3128	0623	- 1151	1374	- 2360	1160
ε_{BB}	-.2624	.0799	-.1843	3574	-.1548	.1758
ε_{FF}	-.6306	.1220	-.5716	2168	-.4961	.2561
ε_{NB}	.1741	.0573	- 0952	.1645	.1188	.1257
ε_{BN}	2041	0528	- 1300	1051	.1291	1005
ε_{NF}	.1386	0315	2102	.0408	1171	.0594
ε_{FN}	4659	1361	.3607	.1903	.4119	.2293
ε_{BF}	.0584	.0265	.3143	.0417	.0257	.0676
ε_{FB}	.1647	0983	2091	.1739	.0842	.2485
Constant-cost input demand elasticities (own and cross)						
e_{NN}	-.9624	.0565	- 8395	.0888	-.9922	.0982
e_{BB}	-.4877	.0733	- 0671	.1441	-.2892	.1503
e_{FF}	-.7429	.1221	- 4088	.2135	-.5980	2555
e_{NB}	- 3717	.0470	- 6954	.1179	-.5755	.0979
e_{BN}	-.0640	.0588	-.2694	.1645	-.0173	1088
e_{NF}	- 0423	.0296	- 0501	.0350	-.0857	0563
e_{FN}	0626	.1372	1851	.1984	.0317	2313
e_{BF}	-.0163	0269	- 0754	0451	- 0136	0670
e_{FB}	-.1741	0997	-.4729	.1835	- 2649	.2477

Contd.

Table : 1 (*Contd.*)

	Pooled Model		Large Farms		Small Farms	
Elasticities	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Allen-Uzawa elasticities						
σ_{NN}	-.6627	1338	-.0280	.2549	-.5159	.2549
σ_{BB}	-.6618	.2072	-.4867	9414	- 3685	4209
σ_{FF}	-4.7956	.9452	-4 2785	2 6723	-4 0437	2 1074
σ_{NB}	4391	.1398	- 2515	4370	.2830	.2954
σ_{NF}	1.0544	.2326	.9755	.4816	.9548	4795
σ_{BF}	4438	.2013	-.6602	5128	.2092	5508
Morishima elasticities						
ρ_{NF}	.7692	1443	.4317	.2417	.6132	2892
ρ_{FN}	7787	1705	5456	2438	.6479	.2850
ρ_{NB}	.4366	.1215	-.2795	.5093	2736	2754
ρ_{BN}	5169	.0623	-.1451	1374	.3650	1160
ρ_{BF}	6889	1320	.2973	.2348	5217	.2981
ρ_{FB}	.4271	.1437	-.3934	4539	.2389	3488
Shadow elasticities						
κ_{NF}	.7712	.1417	.4468	.2176	6205	.2632
κ_{NB}	.4732	.0755	- 2240	2658	.3173	.1478
κ_{BF}	.6237	.1188	.1742	2568	.4578	2886

Note· Subscripts y, c, N, B, F, K, L, M refer to output, cost, hired labour, bullock labour, fertiliser, land, and family labour, respectively.

Appendix : Parameter Estimates

	Pooled Model		Large Farms		Small Farms	
Parameter	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
α_0	3 6499	.0329	4.0661	0446	3.2876	.04479
α_N	-.1972	.0147	- 0873	0198	-.1307	.0167
α_B	-.1240	.0113	-.0508	.0119	-.0893	0122
α_F	- 0385	0037	-.0165	.0040	- 0291	.0042
α_C	.3597	0270	.1546	.0350	.2491	.0316
β_{NN}	.0565	.0096	0178	.0071	.0446	0127

Contd.

Appendix (*Contd.*)

	Pooled Model		Large Farms		Small Farms	
Parameter	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
β_{BB}	-.0730	0100	- 0752	.0172	- 1109	0175
β_{FF}	- 0085	.0045	- 0085	0030	- 0129	.0080
β_{NB}	0589	0080	0374	0094	.0431	.0127
β_{NF}	.0113	0050	0028	.0028	0101	0073
β_{BF}	.0055	0037	0035	0026	.0007	0078
β_{CC}	1262	.0139	.0214	.0101	.0287	0166
ϕ_K	0874	.0255	1309	.0363	0966	0333
ϕ_L	.4153	.0673	7030	1182	8313	1939
ϕ_M	0424	0241	0339	0285	0723	.0380
μ_{KK}	- 0223	.0215	0137	.0305	-.0311	.0293
μ_{LL}	.4726	2289	2981	6008	-.7335	1 7760
μ_{LK}	0092	.0391	0561	.0811	.1017	1236
μ_{MM}	01278	.0151	0027	0153	0256	.0284
μ_{MK}	-.0024	0110	.0027	.0107	- 0091	.0227
μ_{ML}	- 0084	.0291	- 0599	0420	1268	1181
γ_{NK}	0111	0044	0099	.0039	.0144	0055
γ_{NL}	.1486	.0152	- 0028	.0165	2113	0427
γ_{BK}	0006	0053	-.0031	.0028	0051	.0059
γ_{BL}	1110	.0149	-.0423	.0163	- 0408	0363
γ_{FK}	-.0009	0015	- 0033	0011	0019	.0020
γ_{FL}	.0393	.0050	- 0025	.0044	0430	.0143
γ_{NM}	- 0062	0038	0041	0029	0345	.0069
γ_{BM}	- 0041	0033	- 0006	.0019	.0085	0052
γ_{FM}	-.0025	.0011	- 0009	0007	-.0001	.0020
γ_{NC}	-.1267	.0093	-.0581	.0129	- 0978	.0126
γ_{BC}	.0087	.0063	0343	.00819	.0671	.0104
γ_{FC}	-.0082	0022	0023	.0013	.0021	0030
θ_{KC}	- 0109	.0091	- 3588	.0057	-.0214	.0115
θ_{LC}	- 2989	.0302	.0475	0305	- 2136	.0760
θ_{MC}	.0072	.0064	-.0026	0046	- 0430	0112

Economics of Health, Nutrition and Development - Theory and Evidence

Biswajit Chatterjee

If Health is wealth, and wealth of nations and well-being of the people are prime concerns for economic analysis, then Economics of Health should occupy the centre place of economic analysis and discourse. Unfortunately, this is in general not the case with economics profession, and medical economics or health economics is considered to be too narow a domain where the applicability of standard tools of economic analysis is usually viewed with some scepticism. Such a narrow view of the importance of health economics is based on the presumption that allocation of resources for the provision of medicare is insignificant in most of the countries compared to the producing sectors, and however important the R and D investments in medical sciences may be, they do not call for macroecenomic management like the control of money supply, change in the rate of interest and/or sterilization of foreign exchange reserves etc. In recent years however, there have been a revival of a broad.consensus in economic analysis which prefers to give weightage to the nutritional status and the state of morbidity of the people in the discussion on welfare economics and economic development. It is emphasized in recent researches in development economics that health and nutrition are not only important as ends in themselves, but they may constitute channels through which productivity and distributional goals of developing societies may be pursued effectively. Accordingly, in recent years, economists have directed increasing attention towards exploring the determinants of and the impact of health and nutrition in the development process. Important issues in this area concern the extent of health and nutritional problems, the determinants of health and nutrient status, the impact of health and nutrition on incomes and other outcomes, and the choice of appropriate policy designs.

According to the World Development Report, 1993, improved health contributes to economic growth of a nation in four ways: (a) it reduces production losses caused by worker illness; (b) it permits the use of natural resources that had been totally or nearly inaccessible because of disease; (c) it increases the enrollment of children in school and makes them better able to learn, and (d) it frees for alternative uses resources that would otherwise have to be spent on treating illness. While channel (a) signifies the importance of health and nutrition in improving the productivity of workers and their wages - the so-celled efficiency-wage hypothesis–, and also enables the employers to organize skill-formation and specialization through better training, channel (c) corresponds to human capital formation through inter-generational nutritional development, thus affecting the profile of future income streams of the nation in the upward direction. Channels (b) and (d), on the other hand, focus on the efficiency of resource allocation and its acquisition in the current generation so as to step up the productive base of the developing nation in future. It may be noted that the adverse effects of ill health are usually greatest among the poor people, mainly because they are disease-prone more often, and partly because their income depends exclusively on physical labour and they hardly have the savings to cushion the blow. Therefore, from the point of poverty alleviation in developing nations, too, improvement in the conditions of health of the people is also of paramount importance. According to the estimates of the World Development Report 1993, there is strong statistical associations between income growth and health status across time periods for a large sample of countries for a range of model formulations.

In this paper we propose to focus on the interrelationship of health and nutrition in explaining the standard of living of the people in developing countries. The first section will provide a critical evaluation of the basic issues partaining to these relationships. The alternative paradigms with respect to measurements of malnutrition and their variability and their implications for poverty of nations will be examined in this section. Section II will deal with health and nutritional indicators for the Indian economy, on the basis of some existing studies and present our own estimate about the extent of malnutrition in West Bengal based on NSS data on cereal consumption. In section III, we shall summaize the major conclusions of the paper.

Section I

The proximate determinants of an individual's health and nutrition are usually influenced by numerous decisions taken by him given assets, prices and endowments. Since many individuals in LDCs live in consumption units for which the production-consumption separability conditions are not satisfied owing to incomplete markets or because of productivity effects of health and nutrition input consumption, we may consider the choices made by a household-firm decision making unit. The preference function of such a household unit may be summarized as:

(1) $u = u\ (H^i, C^p, C^i, L^i, e^i, S, T), i = 1, \ldots .n.$

Where H^i is the health of household member i, C^i is his consumption, C^p is the consumption of public good, L^i is the leisure time enjoyed by the i^{th} houselold member, e^i is the level of education of household child, S is the number of surving children, and T stands for the taste and preferences and n is the number of individuals in the household. Such an utility function is maximized subject to a production function, given assets and prices. Such a production may relate the health of the ith individual to the consumption and time use of that individual, education of that person and of other persons who make and implement health-related decisions, and the individual and household endowment. This can be represented by .

(2) $H^i = H\ (N^i, C^i, C^p, n, E^i, E^m\ th^i, n^i, \Omega)$

Where N^i is the nutrient intake of the ith individual, E^i & E^m refer to the education of the ith individual and of the members of his family, respectively, th^i is the time devoted by the ith individual on the health care of the family unit including himself, n^i is the individual endowment like the genetic make-up and Ω stand for the overall endowment of the household, and the other variables are as defined before. While these utility and production functions tend to approximate the influence of certain key variables, there may be some other variables which may affect health of an individual in the household unit in a variety of ways. In the same fashion, we may define the mortality (M^i) of the i th individual if his/her health (H^i) falls short of some critical level, i.e.

(3) $M^i = M\ (H^i - H^*), i = 1 \ldots\ldots\ldots. n.$ The nutrient intake of ith person may on the other hand, depend on

(4) $N^i = N(C^i, tf, \Omega)$, where C^i stand for the food intake, tf is the time and skill spent on food preperation, and Ω stand for the overall household endowment/environment. There have been several studies to estimate these relationships (2), (3) and (4) using both cross-section and time series data, but one must be careful in evaluating these estimates. For example, if individuals with a greater unobserved propensity towards illness are relatively sick despite using to a great extent health-related inputs, the failure to control for simultaneously determined health inputs in the estimation of health production function may underestimate the health impact of such inputs. Moreover, since the relation vary across individuals even in the same household and the distribution of inputs among individuals may not be uniform, the use of household averages in the estimates may be misleading. Further, if lagged responses in health to nutrients are not well represented, empirical estimates may understate the health impact of better nutrition.

What is the exact meaning of the term 'nutritional status'? There seems to be a lot of confusion and ambiguity in the literature with respect to this question. Some studies prefer to emphasize the nutrient intakes as cotermineous with nutritional status, while others refer to health outcomes. If there are many non-nutrient determinants of health status such as water supply, individual and household endowments, community facilities and time spent on health care and preventive activities, then the latter use to typify the nutritional status seems misleading; the use of 'nutriment intakes' to represent the nutritional status seems better. Secondly, a measurement problem exists with respect to the use of fixed conversion factors for food with heterogeneous nutrient qualities, as the nutrient elasticities with respect to income, calculated by applying fixed conversion factors to food elasticities may overstate substantially the true nutrient elasticities, since for marginal food expenditures, other attributes like taste, appearance, status value etc. may be much more important than nutrients. A third possible nutrient measurement porblem is that certain critical data are available only on an aggregate level. Most striking in this regard have been the studies by Reutlinger and Selowsky (1976) and Berg (1981), that characterize malnutrition for large groups of people based on a comparison of their average nutrient intakes with average nutrient requirements. Such studies ignore the fact there are uneven distribution of nutrient intakes and requirements across individuals such that there exist considerable variations around the mean standards. Even in microstudies, there may be sufficient intra-household variations in nutrient intakes so that the use of average data lead to misrepresentation.

But the most controversial problem relates to the definition of nutritional norms or requirements and therefore of malnutrition. Such definitions are of extreme importance in assessing the nature and extent of malnutrition in a less developed economy and in analysing the effects of nutrition on the well-being of the people. In the recent literature of development economics, the 'entitlement approach' to standard of living a la' Sen and others, has gained attention, particularly in regard to defining the scope of detailed public action programmes to prevent entitlement failures, and in this context definition and measurement of malnutrition or under-nutrition become integral part of the analysis. Here one may distinguish between two distinct paradigms of nutrient requirements or norms. The first is the 'genetic potential model' in which the human body is self-regulating and self-optimizing, so that the optimal diet depends on inherent genetic characteristics that differ among individuals. Malnutrition, according to this view, can be measured by the shortfall of calorie intakes compared to this average optimal dietary requirement. The second paradigm is the 'individual adaptability model', emphasized among others by Sukhatmee (1977, 1982), Sukhatme and Margen (1978), Srinivasan (1981, 1985) and Seckler (1, 1982), in which individuals adapt to their environment with differences among themselves with respect to response and efficiency of conversion of nutrient intakes into energy, and with substantial interpersonal, intrapersonal and intertemporal variations without any fixed norm or optimal levels. Adaptation includes shorter-run adjustment to variations over time in energy intakes and expenditures around unchanging means (homeostatis), and longer run adjustments to changes in the means through changes in body weight. Within this paradigm, the term 'nutrient requirements' has a much less precise meaning than with the genetic potential paradigm since there does not exist a true optimal. The extent of malnutrition based on the adaptibility hypothesis would be much less compared to the one based on genetic potential paradigm. For example, Sukhatme (1977) estimated that the incidence of poverty in India declines from 50 to 25 percent in urban areas and from 40 to 15 percent in rural areas if allowance is made for variations in individual calorie requirements instead of using a poverty line based on average calorie requirements. We shall turn to the relationship between poverty and nutritional norms later.

Let us focus briefly on the main issues pertaining to the above debate on nutritional standards Firstly, there are significant inter-individual variations in the conversion of nutrients into nutritional achievements, so that the 'requirement' cannot be specified in an individual-independent way. Secondly, it has been emphasized that there can also be sizeable 'intra-individual' variations over time and that a person can maintain a balanced equilibrium by compensating the lower intakes in some periods by higher intakes in others. Accordingly, the identification of all those falling below a calorie norm as being undernourished could possibly exaggerate the number of people with real nutritional deprivation, since intake data are usually based on short-period samples. Thirdly, there may exist multiple equilibria, with the same person possibly achieving different states of balance at different average levels of nutritional intake. Two types of curative adjustments have been suggested: (a) the first operates through body size, with 'small but healthy' people living on relatively low nutritional intakes without experiencing a dimunition of quality of life or functioning ability; (b) the second is the adaptation of nutritional requirements to variations in intakes for a given body size and allowing an unchanged levels of activity. It may be noted that both adjustments work in the direction of compensating nutritional deprivations, and if they are effective, then low levels of food intake may not amount to undernutrition in any consequential sense. But those adjustments and adaptations are likely to be costly in many cases such that deprivation of nutritional efficiency cannot be settled without refering to these costs, which vary across individuals, society and environments, and together define the state of health.

The above discussion throws light on the relationship between poverty and malnutrition. In the traditional head-count measure of poverty, calculation of the poverty line to distinuish between poor and non-poor involves valuation of a basket of consumption goods that may be considered as minimum dietary requirement for a reference man as decided by nutritional experts. The basic issue has been whether it is appropriate to treat the average dietary requirement as equivalent to a minimum nutritional requirement, as the above debate clearly indicates. While for an economist like Dandekar (1981), it is immaterial to delve into the question of nutritional adequancy of the composition of a diet to evaluate the poverty line and the head-count ratio, the adaptationist nutrition - theorist like P.V. Sukhatme had a different perception, who emphasized the scope for inter-personal as well as intra-personal variations in nutritional intakes and their efficiency both at a time and overtime, with possibilities of costless adaptaion depending on age, sex, body weight and height, activity and environment. For analytical purposes, one should, however, distinguish between concepts of poverty and manutrition. While the former connotes to the inability of a person to command a consumption basket that may be deemed as necessary minimum by some criterion, the latter concept refers to the inability of a person to attain nutritional standard by allowing for the possible source of variations in energy intakes between persons and within a person depending upon factors mentioned above. Based on this analytical distinction between poverty and undernutrition, Dandekar (1981) enunciated the following proposition : *all under-nourished may be poor, but all poor may not be undernourished.* This proposition constitutes the basis of all poverty studies in India since the writing of Dandekar on the subject. It must however be noted that while it is one thing to identify the extent of malnutrition by suitably adjusting the cut-off level for the dynamic nutritional process, the real issues of nutritional deprivation should be judged in terms of much broader concepts like capabilities and functionings, as suggested by Sen (1987), and elaborated on by Dreze and Sen (1989). If people are poor by some standard, and if this poverty is due to any malfunctioning of the economy such as unemployment and low productive capacity of the economy, and if the basic reason for undernourishment is the basic inability of human being to command some basket of goods, and inadequate provision of some basic public health facilities like safe drinking water and medicare arrangements, then from the point of view of economic and social analysis and public policy designs, one need to explain why such situations evolve. That is the basic purpose of theoretical exercise linking poverty, malnutrition and deprivation in a dynamic framework. Dasgupta and Roy (1984) provide one such example, although they have treated poverty and malnutrition as the true manifestation of a single malady, namely lack of employment opportunities. It must be emphasized that what endemic undernourishment often leads to is morbidity and loss of efficiency of labour power, and sometimes starvation and hunger in some scale contribute to this propagation of nutritional deprivation, as the experiences of some of the African nations amply demonstrate.

Section II

In this section, we portray the state of health and nutrition in the Indian economy on the basis of existing studies. In the Eighth Five Year Plan, special emphasis was given to the removal of health and nutritional deprivations in different parts of India with the help of specially designed targeted programmes, which so far remained scattered and ill-coordinated. An assessment of the extent of health and nutritional deprivation would be essential to judge the efficacy potentials of these programmes. We shall also compare the all-India scenario with that of the state of West Bengal, which is run by a Leftist Government since 1977 with well-articulated programmes to strengthen the distributive mechanism in the state. The success story of the state of West Bengal in the areas of land reforms, enhancement of agricultural productivity and rural poverty alleviation along with decentralized planning via elected rural panchayets has been emphasized time again in the discussion on the political economy of Indian development process. A close look at health, nutrition, and morbidity in West Bengal will enable us to assess the success or otherwise of the state in this regard. We shall also present our own estimate of the incidence of malnutrition in West Bengal with the help of NSS consumption expenditure data in terms of a simple measure, called the Intake Deficiency Index (IDI).

All-India Scenario*

Let us start with a macro-view. How does India compare with the rest of the world in terms of health-related indicators? Table 1 below compiles the data from the World Development Report, 1993 to present a summary picture. The burden of disease is measured in the units of disability-adjusted life years (DALYs) so that the loss of healthy life can be indicated.

Table : 1

Health Indicators in India and other Demographic Regions, 1990

	World	Sub Saharan Africa	India	China	Other Asian & Islands	Latin American & the Caribbean	Middle Eastern Crescent	Formerly Socialist Economies of Europe	Established Market Economies
1. Population (Millions 1990)	5,267	510	850	1,134	683	444	503	346	798
2 Life Expectancy Birth 1990	65	52	58	69	62	70	61	72	76
3. Child Mortality Rate, 1990	96	175	127	43	97	60	111	22	11
4a Adult Mortality Rate, 1990 (Ages15–59)–Male	234	381	272	201	243	228	228	281	147
4b. Adult Mortality Rate, 1990, Female	169	322	229	150	177	163	174	112	73
5. Death (Millions) 1990	50 0	7 9	9 3	8.9	5 5	3 0	4.4	3.8	7 1
6 Doctors per 1000 population, 1988–92	1.34	0 12	0 41	1.37	0 31	1.25	1.04	4.07	2 52
7. Nurse–doctor ratio 1988–92	1.40	5.1	1.1	0.5	3.0	0 5	1.5	2.2	2 1
8. Hospital Beds per 1000 population, 1985–90	3 6	1 4	0.7	2.6	1 8	2 7	2 9	11.4	8.3
9 Per capita Health Expenditure ($),1990	323	24	21	11	61	105	77	142	1860
10. Health Expenditure as % of GDP 1990									
10a. Total	8 0	4.5	6.0	3 5	4.5	4 0	4 1	3 6	9 2
b Public Sector	4.9	2.5	1.3	2.1	1.8	2.4	2.4	2 5	5 6
c. Private Sector	3.2	2.0	4.7	1.4	2.7	1 6	1 7	1.0	3 6
11 Daly loss due to (%)									
a Communicable diseases 1990	45 8	71.3	50 5	25 3	48.5	42 2	51.0	8 6	9 7
b Non-Communicable Diseases, 1990	42 2	19 4	40 4	58 0	40 1	42 8	36 0	74 8	78.4
c. Injuries 1990	11.9	9 3	9 1	16 7	11 3	15 0	13 0	16.6	11 9
12 Millions of DALYs lost 1990	1362	293	292	201	177	103	144	58	94
13 DALYs per 1000 population 1990	259	575	344	178	260	233	286	168	117

Source World Development Report, 1993.

It is clear from the above table that India's performance with respect to health has not been particularly impressive compared to different demographic regions. Perhaps, India compares favourably with Sub-Saharan Africa in respect to most of the indicators described in the table. This is despite a relatively high percentage of health expenditure in her GDP (6%), where again the role of the public sector happens to be dismal. Only China with a lower per capita health expenditure than India has been successful in achieving a higher life-expectancy, lower child and adult mortality rates, lower size of abolute dealth, higher doctor and hospital beds per 1000 of population and lower DALYs, compared to India. This suggests that despite claims of the policy makers, India has not made comparable progress in the field of health and nutrition as she did in other areas, although compared to the situation 1960s, health indicators improved for India as whole. But our relative ranking in the world has worsened over the years in respect of health related performances. It may be noted that communicable diseases include Tuberculosis, sexually transmitted disease, human immuno-deficiency virus, diarrhea, childhood infections, malaria, respitory infections, maternal causes and perinatal causes etc, while non-communicable diseases include cancer, nutritional deficiencies, neuro-psychiatric disease heart diseases and cerebrovascular diseases. While the incidence of second category (i.e. non-communicable) of diseases are high in formerly socialist economies and established market economies, the incidence of nutritional deficiencies are the lowest in these countries (1.4% & 1.7% respectively), while in India the incidence of nutritional deficiency is the highest at 6.2%, with world average being only 3.9%. Thus one needs to have a close look at the incidence of nutritional deprivation in India during the period of her planned economic development. We shall refer to certain estimates obtained with the help of NSS data.

The head-count measure of under-nutrition based on average calorie requirement of 2780 Kcal for the consumer unit for rural India as a whole was estimated by Dandekar (1981) using NSS calorie intake distribution for the year 1971-72. Ranking households by both energy intake and monthly expenditure per consumer unit, one many estimate the average calorie intakes of various household deciles. Such estimates, due to Kakwani (1989), are presented in table 2 below. Kakwani also estimated (Table-3) alternative head-count measures of undernutrition by assuming the average calorie norm of 2780 Kcal per consumer unit with a coefficient of variation of 15% around this value for both uniform distribution as well as normal distribution of calorie requirements. For alternative values of the coefficient of variations (0-30%) the simulated measures of undernutrition for both rural and urban areas of India during 1971-72, are presented in table 4 below. Certain observations may be made on the basis of these information. Firstly, there exists substantial under-estimation when the estimates of undernutrition are derived from data based on expenditure ranking in stead of ranking by calorie intakes. Secondly, the average calorie norm approach provides estimates which are very close to those obtained by using the entire distribution of calorie intakes. Moreover, since the estimates of undernutrition obtained by using uniform and normal distributions are almost identical, perhaps the choice of the distribution of calorie requirement does not make much difference. Finally, it is clear that the estimates of undernutrition is not very much sensitive to variations in coefficient of variations. What is perhaps more important is that when the calorie intake and the calorie requirements are correlated, the degree of undernutrition monotonically diminishes with the correlation coefficient in rural areas, but in urban areas it shows a monotonic increase. This is evident from the simulations of Kakwani (1989), presented in Table 5.

Table : 2

Average calorie intake of various household deciles when households are ranked according to energy intake per diem per consumer unit and monthly expenditure per consumer unit in all Indian Rural areas, 1971-72

Household Deciles	Average Calorie intake ranked by energy intake per diem per consumer unit	Average calorie intake ranked by monthly expenditure per consumer
1	1333	1675
2	1818	2094
3	2093	2331
4	2341	2541
5	2586	2745
6	2849	2960
7	3145	3197
8	3508	3482
9	4040	3880
10	5843	5058

Source Kakwani (1989)

Table : 3

Head-count Measures of undernutrition based on the distribution of individuals - All India, 1971-72

Head-Count Measure	Average calorie norm approach	Sukhatme approach	Uniform distribution	Normal distribution
1. Rural Areas	52.4	18.8	51.3	51.4
2. Urban Areas	67.4	23.0	64.2	64.5
3. Ratio of Rural to Urban undernutrition	0.776	0.817	0.799	0.797

Source Kakwani (1989)

Table : 4

Head-count Measures of undernutrition in India during 1971-72 for alternative values of coefficient of variations.

Coefficient of Variation	RURAL		URBAN	
	Uniform Distribution	Normal Distribution	Uniform Distribution	Normal Distribution
0.00	52.4	52.4	67.5	67.5
2.50	52.4	52.4	67.4	67.4
5.00	52.3	52.3	67.1	67.1
7.50	52.1	52.1	66.6	66.6
10.00	51.9	51.9	66.0	66.0
12.50	51.6	51.6	65.2	65.3
15.00	51.3	51.4	64.2	64.5
17.50	50.9	51.1	63.3	63.7
20.00	50.2	50.5	61.2	62.1
22.50	50.2	50.5	61.2	62.1
25.00	49.8	50.3	60.3	61.3
27.50	49.5	50.0	59.4	60.6
30.00	49.2	49.8	58.6	59.7

Source : Kakwani (1989)

Table : 5

Head-count Measures of undernutrition in India during 1971-72 with alternative correlation Coefficients between intake and requirement of calories.

Coefficient of correlation	Rural Areas	Urban Areas
0.00	51.40	64.50
0.10	51.07	64.64
0.20	50.79	64.86
0.30	50.50	65.16
0.40	50.21	65.56
0.50	49.93	66.09
0.60	49.65	66.79
0.70	49.37	67.73
0.80	49.12	69.02
0.90	48.90	70.88

Source : Kawani (1989)

The above discussion points out the fact that the head-count measure of undernutrition, however rough and ready it might be in terms of methological accuracy, remains more or less insensitive to alternative values of coefficient of variation in calorie intakes around the average requirement norm, and the correlation coefficient between intake requirements and actual intakes of a hypothetical consumer unit. The order of estimation biases remains very small when the average requirement norm is used as a cut-off level to estimate undernutrition. Therefore, the head-count measure of malnutrition gives us on an average the extent of nutritional deprivation in our country, which is however quite high. But has there been any marked improvement since the 1970s ? Unfortunately, the answer is in the negative. Table 6, which presents the actual intake data collected by the National Nutrition Monitoring Bureau for the period 1975-82 for the rural areas of nine Indian states indicate that not only are average levels of calorie intake lower than the average intake requirements in most states, but also that there is no sustained increase in calorific intakes during the period. Thus for the country as whole, whether one goes by calorie based or direct anthropometric measures, about 50% of the rural population is undernourished; hence the urgency of designing appropriate mode of public intervention in this regard.

Table : 6

Average Calorie Intake in Rural India during 1975-82

States	1975	1976	1977	1978	1979	1980	1981	1982
1. Kerala	1,926	2,052	1,722	1,805	2,019	2,158	-	2,203
2 Tamilnadu	-	2,249	2,477	2,411	2,537	2,196	2,346	1,964
3. Karnataka	2,911	3,058	2,588	3,008	2,751	2,992	2,873	2,711
4. Andhra Pradesh	2,394	2,569	2,665	2,527	2,600	2,391	2,238	2,061
5. Maharastra	2,203	2,315	2,407	2,295	2,282	-	2,472	2,120
6. Gujrat	2,203	2,146	2,177	2,131	2,327	2,333	2,162	2,306
7 Madhya-Pradesh	-	2,393	2,045	2,180	2,205	-	-	-
8 West-Bengal	2,144	2,473	2,381	2,584	2,177	2,580	2,477	2,426
9. Uttar-Pradesh	2,283	2,064	2,292	1,956	1,983	2,115	2,193	-
10. Average	2,296	2,369	2,306	2,341	2,366	2,404	2,408	2,243

Source NNMB data quoted in Kunar & Stewart (1992)

The NSS consumer expenditure survey data for the 38th Round (1983) covered quantity estimates of each of the food items consumed by each of the households in the sample of 123,000 spread over rural and urban areas in India. These quantity estimates were converted into their equivalent intakes of calories with the help of appropriate conversion factors. After adjustment for free meals obtained by households and the loss of food component at the time of preparation, B. S Minhas (1991) has estimated that in 1983, 37.4% of the rural and 43.71% of the urban households remain below the 2,430 caloric level (90% of the norm). The corresponding percentage below the 80% level i.e. 2,160 calories are 35.09 and 29.29 respectively, in rural and urban India.

The distributions of households over five broad groups of per consumer unit per diem calorie intake after adjusting reported food consumption for free meals are given in table 7 for different states in India. The proportion of households having meal-adjusted calorie intake of less than 2,160 K cal varies in the rural sector from 11-15% to 38, 39 and 43 per cent, respectively in West Bengal, Tamil Nadu and Kerala, the average proportion for rural India as a whole being 25%. Approximately, 15% of the rural households in India had energy intakes of less than 1, 900 Kcal. West Bengal (24.5%), Tamil Nadu (26.8%) and Kerala (29%) reported very high proportions of rural households where actual calorie intake were even less than 70% of the recommended nutritional norm of 2,700 K cals (ie 1900 K cals). Similar picture also emerges for the urban households in 1983. Thus we not only have significant undernutrition both in rural and urban sectors in India during 1983, but there exists also considerable inter-state diversity with respect to such nutritional deprivations.

Table : 7

Percentage Distribution of Households over Five Broad Groups of Per Consumer Unit Per Diem Caloric Intake After Adjusting Reported consumption for Meals : States and All-India, 1983

States	Calories < 1890		Calories < 2160		Calories < 2430		Calories within 2430-2700		Calories above 2700	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
1 Andhra Pradesh	11.80	17.22	21.18	30.36	34.66	46.38	13.92	14.40	51.42	38.77
2. Assam	12.62	15.27	31.28	27.70	44.91	41.48	16.87	19.74	38.22	38.78
3. Bihar	14.41	11.87	24.93	21.78	37.36	35.48	13.27	15.48	49.37	49.04
4 Gujrat	16.62	18.66	29.11	32.11	43.20	45.76	15.34	17.18	41.46	37.06
5 Haryana	9.20	11.60	14.62	21.17	23.02	34.65	12.67	11.73	64.31	53.62
6. Himachal Pradesh	6.33	7.97	11.31	16.00	17.75	19.38	9.81	19.12	72.44	61.50
7. J. & K	5.73	6.88	11.09	14.66	18.13	32.33	11.21	17.36	70.66	50.81
8 Karnataka	18.48	21.62	28.42	32.35	40.07	42.94	11.13	11.89	48.80	45.17
9 Kerala	29.00	26.48	43.05	37.83	55.48	50.27	11.07	9.87	33.45	39.86
10. Madhya Pradesh	9.18	10.47	17.59	20.86	29.89	36.48	13.35	16.83	56.76	46.69
11 Maharastra	13.35	16.88	25.84	31.88	40.22	48.37	14.37	15.15	45.41	36.48
12 Orissa	18.24	9.03	29.99	18.16	43.31	30.37	15.58	20.09	44.11	49.54
13 Punjab	8.81	18.06	15.08	28.88	22.70	40.20	10.79	12.35	66.51	47.54
14 Rajasthan	10.61	12.04	16.28	21.00	25.41	34.98	10.74	12.96	63.85	52.06
15 Tamil Nadu	26.83	28.14	39.00	42.51	51.77	55.53	11.02	12.00	37.21	32.39
16 Uttar Pradesh	9.43	14.51	17.35	27.23	28.15	42.91	11.95	14.11	59.90	42.98
17 West Bengal	24.47	14.50	37.69	27.11	51.15	42.23	13.35	17.19	35.32	40.58
18. All-India	14.58	17.03	25.09	29.29	37.40	43.71	12.77	14.58	49.83	41.71

Source · NSS 38th Round Draft Report No. 356, 1989, quoted in Minhas (1991)

West Bengal Scenario

In order to estimate the incidence of malnutrition among people of different expenditure classes in West Bengal, we have used the NSS survey data on malnutrition in 27th and 38th rounds. First of all, we compute the percentage of population below the poverty line (ie head-count ratio measure of poverty) for both rural and urban areas of West Bengal on the basis of poverty lines estimated from average Bengali diet recommended by nutritional experts and using retail consumer prices. Thus the state-specific diet and state-specific prices are used to estimate the poverty line. Secondly, to estimate the incidence of undernutrition, we take the recommended minimum calorie intake of a 'reference man' engaged in moderate activity and calculate Sukhatme's modified head-count measure of undernutrition, namely, (C-3σ), where σ stand for the standard deviations of inter-individual and intra-individual variations in nutritional requirement depending upon height, body weight, metabolic absorption, type of activities engaged, age, environment, etc, and C is the recommended minimum calorie intake of the reference man per day. Actual calorie intakes falling short of (C-3σ) will constitute nutritional deprivation as per this measure, (C-3σ) being estimated to be equal to 1540 K cal per day per person. Similarly to estimate the protein deficiency we take the cut-off intake level of 55 grams per day per person as the absolute minimum requirement. With the help of these all India intake requirements, we calculate the extent of calorie and protein deficiency in rural and urban areas of West Bengal and compare our estimates with the corresponding all-India estimates. This is given in table 8 below :–

It may be noted that for all expenditure classes, there are significant correlations between mean monthly per capita consumption expenditures and the average intake of calories and the average protion intakes such that although poverty and undernutrition are not synonimous, there exist significant associations between the two. In fact, one of the reasons for nutritional deprivation happens to be the inability of various expenditure groups to afford and attain the required norms as suggested by nutritional experts. We have estimated an Intake Deficiency Index (IDI) which measures the percentage deviations of actual consumption from the recommended norm of average Bengali diet with respect to cereals for different monthly expenditure classes in rural and urban areas of West Bengal for the different NSS rounds - 28th, 32nd, 38th and 42nd. We have used the recommended intake of 550 grams and 450 grams of cereals per capita per day as the minimum cereal diet for average Bengali workers in rural and urban areas of West Bengal as per the recommendations of the All India Institute of Hygiene and Public Health, Calcutta. The rationale for concentrating on cereal consumption only is simly that the NSS consumer expenditure data give direct quantity estimates of cereals and cereal substitutes only, and that it is difficult to obtain price data from the NSSO to deflate expenditures on other components of food consumption We have accordingly, chosen to ignore the substitution possibilities between alternative items of consumption of food in average Bengali diet and measure the IDI with respect to cereals only, which constitute almost 60% of the expenditure on food in an average Bengali diet. It is interesting to note that positive value of the Intake Deficiency Index implies nutritional deprivations and when the index takes negative value, there is excess nutrition and hence waste. Secondly, there is a basic asymmetry in the paterrn of cereal consumption between people in rural and urban areas of West Bengal, respectively. People in the urban areas for all expenditure classes have intake deficiency of cereals because they in general consume less cereals than their rural counterparts and consume non-cereals more and they also allocate a large portion of their total expenditure on non-food items like education, shelter, health etc Thirdly, the incidence of intake deficiencies are usually found to be larger for lower expenditure classes in both rural and urban areas, and on an average, taking all expenditure classes together, we have positive IDIs or nutritional deprivations in all NSS round in both rural and urban areas, although the IDIs themselves have shown signs of going down over the successive rounds. To assess the interdependence between poverty and nutritional deprivations in the state, we have estimated the average IDI (%) below the poverty lines for both rural and urban areas in the state, and the poverty lines using state-specific Bengali diet and state-specific retail prices, are found to be Rs.62.95, Rs. 72.29, Rs.124.51 and Rs.139.14 for rural West Bengal in the four NSS rounds respectively, while the corresponding urban poverty lines are Rs.69.51, Rs.90.00, Rs.129.00 and Rs 175.51 respectively, all in the units of per capita per month. Also we have estimated the Poverty gap (PG) Ratio (%) measure of poverty and the correlation coefficient between PG ratio and the average intake deficiency index below the poverty line. These estimates are given in table 9 below :-

Table : 9

Average Intake Deficiency Index (%) below the Poverty Lines in West Bengal–Rural and Urban, and the Poverty Gap Ratio (%) in various NSS Rounds.

	RURAL			URBAN		
NSS Rounds	Average Intake Deficiency Index (%) below the Poverty Lines	Percentage Variations over the preceding Round	Poverty Gap Ratio (%)	Average Intake Deficiency Index (%) below the Poverty Lines	Percentage Variations over the preceding Round	Poverty Gap Ratio (%)
1. 28th Round (1973-74)	29.75	-37.38	57.2	29.43	-53.89	23.2
2. 32nd Round (1977-78)	18.63	20.07	49.7	13.57	32.20	31.5
3. 38th Round (1983)	22.37	-27.04	45.3	17.94	-22.02	22.6
4. 42nd Round (1986-87)	16.32		34.3	13.99		20.9
5. Coefficient of Variations (%)	46.76		35.74	68.39		33.40
6. Correlation Coefficient with P.G Ratio (%)	0.83			-0.31		

Source : Computed from NSS consumer expenditure data, various rounds

Certain observations may emerge from the above estimates. First of all, the degree of variations in the average intake deficiency index below the poverty lines for both rural and urban areas in West Bengal much exceeds the corresponding variations in poverty gap ratios across the various rounds. This suggests that the intensity of nutritional deprivation among the poor in West Bengal has been larger than the intensity of poverty itself. Secondly, we may argue that although for rural areas in West Bengal poverty and under-nutrition are significantly correlated with the expected sign, for the urban areas, the degree of association is much weaker and further that improving the nutritional efficiency with respect of cereal consumption does not perhaps contribute to the reduction in the intensity of poverty among the urban poor in West Bengal, and other socio-economic variables appear to be crucial in minimizing the intensity of urban poverty in the state.

How does West Bengal compare with all India averages with respect to medical facilities across various expenditure classes ? The NSS 42nd round survey on Morbidity and Utilization of Medical Services, points out an interesting indirect evidence in this respect. The report classifies people in rural and urban areas belonging to various income/expenditure levels into fractile groups and provides the percentage distribution of untreated spells of sickenss by reasons for non-treatment in India and West Bengal for both rural and urban areas separately. We have picked up three main reasons for non-treatment, namely; Absence of Medical Facility, Long Waiting which indicates the insufficiency of medical facility, and Financial Reasons, which is indicative of poverty. Interestingly enough, Zero medical facility causing non-treatment is serious among the bottom three decile groups for the all-India rural scenario, whereas for West Bengal, this factor is not at all a contributing factor for those three bottom decile groups in rural areas. No medical facility appears to be most important in rural West Bengal among the fractile groups 70-80. Long-waiting is not seen to be an important factor in rural West Bengal, whereas it is important for India as a whole among almost all fractile groups. Financial reasons or poverty happens to be a very important factor for both West Bengal and all-India, although in respect of rural areas, West Bengal has a slightly better score than all-India figures. For urban areas, on the other hand, financial reasons appear to be the most important factor for non-treatment of illness in West Bengal among the bottom *two* decile classes, while in India as a whole, the importance of this factor is distributed among the bottom *three* decile groups. No medical facility happens to be a non-factor for non-treatment of illness for most of the fractile groups of urban West Bengal, while long-waiting happens to be an important cause among the fractile groups, 0-10, and 80-90 only. This suggests that in West Bengal there has been some progress in the extension of medical facility among the poorer sections of the population which is relatively satisfactory vis-a-vis the all India situation both with respect to the existence of facility and its sufficiency. However, financial reasons continue to be important for both rural and urban areas in West Bengal, which indicates that given the escalation of costs of medical treatment over the years, poor people have to tolerate with sickness and morbidity thereby affecting not only their quality of life in general, but also their ability to combat malnutrition, hunger and poverty, which is reflected in low labour productivity and low wage rates among vulnerable section of the population, particularly in rural areas.

Table : 10A

Percentage Distribution of Untreated Spells of Sickness over Fractile Groups by Reasons for Non-Treatment in India and West Bengal-Rural areas, 1986-87

	REASONS FOR NON-TREATMENT					
	All-India			West Bengal		
Fractile Groups	No Medical Facility	Long Waiting	Financial Reasons	No Medical Facility	Long Waiting	Financial Reasons
0-10	18 84	11 51	21 75	-	-	15 69
10-20	10 70	12 01	13 20	-	-	15.61
20-30	9.00	5.89	12.68	-	-	10 89
30-40	10.23	17 15	11 23	12 25	-	11 93
40-50	6.00	4 55	8 45	8.58	-	13 63
50-60	7.19	4 81	8.31	9.57	-	8 58
60-70	15.03	14 24	8 32	16.12	-	9.32
70-80	13 14	13 53	7 89	47 84	-	4.24
80-90	7.94	9.93	6.80	43.86	-	6.93
90-100	1.93	6 38	1 31	1.78	-	3 18
All	100.00	100 00	100 00	100 00	-	100 00

Table : 10B

Percentage Distribution of Untreated Spells of Sickness over Fractile Groups by Reasons for Non-Treatment of India and West Bengal - Urban areas.

	REASONS FOR NON-TREATMENT					
	All-India			West Bengal		
Fractile Groups	No Medical Facility	Long Waiting	Financial Reasons	No Medical Facility	Long Waiting	Financial Reasons
0-10	18 46	20 82	29.82	-	22.77	40 87
10-20	-	22.59	18 75	100 00	-	44 96
20-30	31.21	12 30	24 23	-	-	3 02
30-40	4 06	7 83	6.07	-	-	-
40-50	-	3 85	5 91	-	-	8 40
50-60	-	1 98	6 66	-	-	0 71
60-70	-	3 09	4 64	-	-	-
70-80	40 18	1.68	2.47	-	-	-
80-90	-	25.86	1 37	-	77 23	2.04
90-100	6 09	-	0 03	-	-	-
All	100.00	100.00	100 00	100 00	100 00	100.00

Source of 10A + B NSS Report on Morbidity and Utilization of Medical Services, 42nd Round, July 86 - June 1987, No 364, Table 12

Section III

It may be mentioned that in our discussion on the relation between health and nutrition in the context of India and West Bengal we have not touched upon the role of the state intervention, and instead concentrated on the estimates of nutritional deprivations. This is primarily because of two reasons : (a) Knowledge of the extent of deprivation is essential in defining the scope of public action, and (b) the choices involved in devising appropriate public health and intervention programmes are varied and much too complex to be discussed within the ambit of one single paper. In a LDC like ours,removal of endemic hunger and nutritional deprivation involves hugh public investment on different aspects like poverty alleviation, food production, procurement and distribution, expenditure on medical facilities and infra-structure, education, apart from employment generation through a process of rapid economic growth. Overall macroeconomic management of the economy and the design of 'structural adjustment programmes' are intimately related to these issues on public action on health and nutrition. Our study emphasising the theoretical relationships between health and nutrition and measures of nutritional deprivation and morbidity in India and West Bengal, focuses on a curcial aspect of economic development which has been ignored in the early years of our planning. The recognition of these crucial relationships in the literature on development economics in recent years has prompted shift of focus in policy designing from pure commodity fetishism to the domain of capability and functioning of human being in the process of economic development of a nation. This paper has attempted an evaluation of nutritional well-being in India and West Bengal from the point of view of this 'capability' paradigm.

In particular, certain specific conclusions emerge from our statistical estimates. They may be summarized as follows :

(a) All undernourished may be poor, but all poor may not be necessarily undernourished;

(b) The extent of nutritional deprivations in India and the states are quite substantial and there exist considerable inter-state variations in the measures of nutritional intake as well as in their deficiencies;

(c) The head-count measure of malnutrition, however deficient might be with respect to variability in calorie intakes between and within persons,are quite robust in response to alternative specifications of coeficient of variations and correlation coefficient to take account of the required adjustments;

(d) The extent of nutritional deprivations in rural and urban West Bengal is found to be quite significant despite a declining trend, and financial reasons or poverty continues to be a major factor for the morbidity of the prople in the state;

(e) There exists a high degree of association between the average intake deficiency below the proverty line and the intensity of poverty as measured by the poverty gap ratio for rural areas in West Bengal, but for the urban areas, the relationship is found to be inverse, but weak; and

(f) India, as a whole, does not compare favourably with rest of the world in respect of the indicators of health, and that there exists considerable scope for public action in this regard in our country during the coming years.

A final world of concem may be noted at this stage. While our paper points to the need for widening the scope of public intervention programmes in health and nutrition in India and West Bengal and for targeting these programmes effectively towards the under previledged, the abilities of the Govt. of India and of the Govt. of West Bengal in launching such huge public investment expenditures in the near future appear limited in view of the severe financial crunch prevailing in the economy and the structural adjustment programmes put into practice by the Govemment of India in recent years. The basic tenor of this structural adjustment programme has been to reduce the size of govemment expenditure and with it the role of the state, and the sectors which are likely to be hard hit are activities like education, health, sanitation, drinking water and other social services and infrastructure etc. Although the World Bank has emphasized the need for allocating more funds towards improvement of health,nutrition and environment in India, this is unlikely to be achieved in near future unless concessional loans and grants from multilatural agencies are forthcoming in large amounts, and even so, unless the government clearly formulates well-targeted plans to eradicate this endemic deprivations among millions of people. But as our experience of economic development demonstrates, in India political rhetoric often dominates over sound economic rationale and maximization of social welfare. Consequently, deprivations remain.

REFERENCES

1. Behrman, J R. and Deolalikar, Anil, B. (1988)- 'Health and Nutrition', Chapter 14 of *Handbook of Development* Economics, Vol I, edited by H.B. Chenery and T N Srinivasan, North-Holland Publishing Company, 1988.

2 Berg A.D. (1981)- 'Malnourished People · A Policy View' World Bank, mimeo 1981

3 Chatterjee Biswajit and others (1991)- *Incidence of Poverty and Poverty Alleviation Measures in West Bengal,* World Bank, 1991.

4 Dandekar, V.M (1981) 'On the Measurement of Poverty', in Economic and Political Weekly, July 25, 1981, reprinted in Krishnaswamy K.S. (ed)- *Poverty and Income Distribution,* O'xford University Press, 1990

5 Dasgupta, P and Ray, D. (1984)- "An Economic Theory of Malnutrition', in Guhan S and Shroff M (eds)- *Essays on Economic Progress Welfare, in honour of I.G Patel,* Oxford University Press, 1984.

6. Dasgupta, P and Ray, D (1990)- 'Adapting to undernourishment The Biological Evidence and Its Implications', In Dreze, J and Sen A K (ed) - *The Political Economy of Hunger,* Vol I, Clarendon Press, 1990.

7 Dreze J and Sen, A K - *Hunger and Public Action, Clarendon Press, 1989.*

8 Kakwani N. (1989) - 'On Measuring Undernutrition' in *Oxford Economic Papers,* 1989, p.p. 528-552.

9 Kumar, Gopalakrishna and Stewart, F. (1992) " "Tackling Malnutrition . What can Targeted Nutritional Interventions Achieve ?" in Barbara Harriss (ed) - *Poverty in India : Research and Policy,* Oxford University Press, 1992.

10. Minhas B.S. (1991)- 'On Estimating The Inadequecy of Energy Intakes · Revealed Food Consumption Behaviour versus Nutritional Norms (Nutritional Status of Indian people in (1983) in *Journal of Development Studies,* October, 1991, pp 1-38.

11. National Sample Survey Organizations - Reports on consumer Expenditure, various rounds, and on Morbidity and Utilization of Medical Services, 42nd Round.

12 Panchamukhi. P.R (1992)- *Economics of Health and Nutrition,* Centre for Multi-disciplinary Development Research Dharwar, 1992.

13. Reutlinger, S. and Selowsky M (1976)- *Malnutrition and Poverty : Magnitude and Policy Options,* Johns Hopkins University Press.

14 Seckler, D (1982)- 'Small but Healthy A Basic Hypothesis In the Theory, Measurement and Policy of Malnutrition', in P.V Sukhatoc, (eds)- *Newer concepts in Nutrition and their Implications for Policy, Maharashtra Association for the Cultivation of Science Research Institute, Pune*

15 Sen, Amartya K. (1987) . *The Standard of Living*, Cambridge University Press, 1987.

16. Srinivasan, T N. (1981) . 'Malnutrition : Some Measurement and Policy Issues', *Journal of Development Economics, 1981*, pp 3-19

17 -do- (1985) · '*Malnutrition in Developing Countries · The State of Knowledge*', mimeo

18 Sukhatme, P V (1977) 'Malnutrition and Poverty', 9th Lal Bahadur Shastri Lecture, New Delhi.

19. -do- *ed* (1982) - *Newer concepts in Nutrition and their Implications For Poverty*, Maharastra Association For the Cultivation of Sciences Research Institute 1982.

20 Sukhatme P.V and Margen. S (1978) . 'Models For Protein Deficiency' in *American Journal of Clinical Nutrition.*

21. World Bank (1993) - *The World Development Report, 1993.*